UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Tecnologia

JOSÉ RENATO PAVIOTTI

CONSIDERAÇÕES SOBRE O CONCEITO DE ENTROPIA NA TEORIA DA INFORMAÇÃO

LIMEIRA
2019

JOSÉ RENATO PAVIOTTI

CONSIDERAÇÕES SOBRE O CONCEITO DE ENTROPIA NA TEORIA DA INFORMAÇÃO

Dissertação apresentada à Faculdade de Tecnologia da Universidade Estadual de Campinas como parte dos requisitos para a obtenção do título de Mestre em Tecnologia, na área de Sistemas de Informação e Comunicação.

Orientador: Prof. Dr. José Carlos Magossi

ESTE TRABALHO CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELO ALUNO JOSÉ RENATO PAVIOTTI, E ORIENTADA PELO PROF. DR. JOSÉ CARLOS MAGOSSI.

LIMEIRA
2019

Agência(s) de fomento e nº(s) de processo(s): Não se aplica.

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Faculdade de Tecnologia
Felipe de Souza Bueno - CRB 8/8577

P288c

Paviotti, José Renato, 1985-
Considerações sobre o conceito de entropia na teoria da informação / José Renato Paviotti. – Limeira, SP : [s.n.], 2019.

Orientador: José Carlos Magossi.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Tecnologia.

1. Comunicação. 2. Matemática. 3. Tecnologia. 4. Entropia (Teoria da informação). I. Magossi, José Carlos, 1963-. II. Universidade Estadual de Campinas. Faculdade de Tecnologia. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Considerations about the concept of entropy in the information theory
Palavras-chave em inglês:
Communication
Mathematics
Technology
Entropy (Information theory)
Área de concentração: Sistemas de Informação e Comunicação
Titulação: Mestre em Tecnologia
Banca examinadora:
José Carlos Magossi [Orientador]
Renato Kraide Soffner
Werner Martins Vieira
Data de defesa: 27-02-2019
Programa de Pós-Graduação: Tecnologia

FOLHA DE APROVAÇÃO

Abaixo se apresentam os membros da comissão julgadora da sessão pública de defesa de dissertação para o Título de Mestre em Tecnologia na área de concentração de Sistemas de Informação e Comunicação, a que submeteu o aluno José Renato Paviotti, em 27 de fevereiro de 2019 na Faculdade de Tecnologia – FT/UNICAMP, em Limeira/SP.

Prof. Dr. José Carlos Magossi

Presidente da Comissão Julgadora

Prof. Dr. Renato Kraide Soffner

UNISAL e FATEC

Prof. Dr. Werner Martins Vieira

UNISAL

Ata da defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós Graduação da FT.

Agradecimentos

Uma das grandes fragilidades das pessoas talvez seja a prepotência ao acreditar que conseguem chegar a algum lugar sozinha. Está completamente equivocada. Chegamos e deixamos este mundo através das mãos de outras pessoas. Por isso, o espaço de apenas uma página será pequeno para reconhecermos e agradecermos as necessidades divinas e humanas que temos.

Por isso, primeiramente agradeço a Deus. Não só pelo dom da vida, mas também pela sabedoria por discernir e escolher os caminhos corretos que possibilitaram a construção de minha carreira, minha família, minha vida. Sabedoria esta construída e fundamentada pelo amor, honestidade e ética. Valores estes herdados e sempre balizados pelo apoio irrestrito e incondicional de meus pais José Carlos e Maria Olívia, aos quais agradeço imensamente.

Agradeço também a minha esposa Kelly por, nestes nove anos, ter participado comigo dos meus sonhos e projetos e ter acreditado que juntos poderíamos vencer quaisquer desafios. Obrigado pela paciência e perseverança, mesmo nos momentos mais difíceis e de pouca esperança. Paciência também demonstrada como uma grande virtude da minha filha Ana Maria que, com apenas cinco anos de idade, sempre nos diz que "esperar é difícil", porém nos surpreendeu ao entender a prioridade de estudar e crescer de seu papai em muitos momentos. Estendo o agradecimento a toda a família: irmãos, sogro, sogra, cunhado, tios e primos por acreditarem, me apoiarem e por entenderem que às vezes temos que nos distanciar para mantermos o foco em alguns objetivos, como este.

Também sou muito grato ao prof. Dr. José Carlos Magossi por acreditar em meu potencial, pela confiança e companheirismo durante todo o período que trabalhamos juntos. Este trabalho é o reflexo e resultado da sinergia produzida por esta interação que sempre foi pautada por valores comuns de respeito, honestidade, ética e muito trabalho nestes pouco mais de dois anos.

Registro também um agradecimento pelas contribuições recebidas por esta pesquisa, realizadas pela banca de qualificação composta pelos professores: Dr. Renato Kraide Soffner, Dra. Talía Simões dos Santos Ximenes e Dr. Werner Martins Vieira. Ao prof. Dr. Werner, um agradecimento especial por ter participado e contribuído no exame de qualificação mesmo na condição de suplente.

Finalizando, agradeço ao Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP), do qual orgulhosamente faço parte do quadro de servidores, pelos incentivos recebidos através da Política de Autocapacitação e do Incetivo Educacional. Ao mencionar o IFSP, também agradeço aos colegas servidores do Câmpus Capivari que torceram e me apoiaram. Em especial, aos colegas de trabalho e grandes amigos Júnio, Larissa Carvalho, Larissa Gatti, Lucas e Washington pela paciência, ajuda e por comporem as "pré-bancas" realizadas informalmente na sala de estar da minha casa, como proposta de ensaiar e receber críticas construtivas que melhoraram consideravelmente a qualidade deste trabalho.

Resumo

Claude E. Shannon, em seu artigo de 1948, lança a pedra fundamental para o que hoje em dia se chama de *Teoria da Informação*. Seu trabalho, motivado por problemas de comunicação entre máquinas, e alicerçado pela matemática, acaba por servir de base para o que hoje se denomina de *era digital*, ou era da informação. Os impactos de sua teoria auxiliaram o surgimento das *Tecnologias de Informação e Comunicação* bem como das ramificações a outros campos científicos que, desde então, crescem e multiplicam-se em ritmo acelerado. A simples ocorrência da palavra *comunicação* faz com que essa teoria seja extrapolada para setores distintos daqueles propostos por Shannon. Com o objetivo de clarear essa situação investiga-se um conceito, na teoria de Shannon, que acredita-se ser aquele que vai possibilitar essa diversidade de interpretações. Esse conceito, defende-se, é o conceito de entropia. É evidente que um conceito com tal plasticidade pode gerar leituras equivocadas, se um cuidado excessivo não for despendido quando da sua utilização. Objetiva-se inicialmente expor algumas considerações históricas acerca do conceito de entropia na Teoria da Informação e de outros termos os quais ora ou outra podem gerar dúvidas em suas leituras. Objetiva-se também indicar o quanto esse conceito é significativo para o desenvolvimento de novas tecnologias. Em particular expõe-se um problema interessante relacionado aos canais de difusão (*Broadcast Channel*) em Teoria da Informação.

Palavras-chave: Comunicação, Matemática, Tecnologia, Shannon, Confusões, *Broadcast*, *Channel*.

Abstract

The Shannon's seminal paper of 1948 is the cornerstone to what nowadays is called *information theory*. His work, motivated by problems in communication between machines, and based on mathematics, turns out to be the basis for what today is called the *digital age*, or the age of information. The effects of his theory aided the rise of *information and communication technologies* as well as ramifications to other scientific fields, which since then, has increase and multiplied at an accelerated rate. The mere occurrence of the word *communication* allows this theory extrapolated to others sectors different from those within the meaning of Shannon. In order to clarify this situation we investigated a concept, in the Shannon theory, which we believe to be the one that allows the diversity of interpretations. This concept is the concept of entropy. It is clear that a concept with such plasticity may produce erroneous readings, if an excessive care is not taken about his utilization. First of all, the objective of this dissertation is to expose some historical considerations about the concept of entropy in information theory and also of other terms, which, from time to time, may generate doubts in his readings. The goal is also indicate how this concept can be significant for the development of new technologies. In particular we expose an interesting problem related to *Broadcast Channel* in information theory.

Keywords: Communication, Mathematics, Technology, Shannon, Confusions, Broadcast, Channel.

Lista de Figuras

Lista de Tabelas

Lista de Abreviações e Siglas

ASC	*American Standard Code for Information Interchange*
BC	*Broadcast Channel*
Bit	*Binary Digit*
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CD	*Compact Disc*
CDMA	*Code Division Multiple Access*
CPF	Cadastro de Pessoas Físicas
DM	*Discrete Memoryless*
DVD	*Digital Versatile Disc*
FFT	*Fast Fourier Transform*
FAX	Facsimile - Equipamento utilizado para fazer telecópias
GB	*GigaByte* - Unidade de Medida equivalente a 1024 *Megabytes*
HDTV	*High-Definition Television*
IEEE	*Institute of Electrical and Electronics Engineers*
ISBN	*International Standard Book Number*
JPEG	*Joint Photographic Experts Group*
KB	*KiloByte* - Unidade de Medida equivalente a 1024 *Bytes*
MB	*MegaByte* - Unidade de Medida equivalente a 1024 *Kilobytes*
MIT	*Massachusetts Institute of Technology*
ml	Mililitro, equivalente a 10^{-3} litros
MPEG	*Moving Pictures Experts Group*
NASA	*National Aeronautics and Space Administration*
PCM	*Pulse Code Modulation*
PNG	*Portable Network Graphics*
SI	Sistema Internacional de Medidas
SMI	*Shannon's Measure of Information*
TIC	Tecnologia de Informação e Comunicação
TV	*Television*
UPC	*Universal Product Code*

Sumário

Capítulo 1

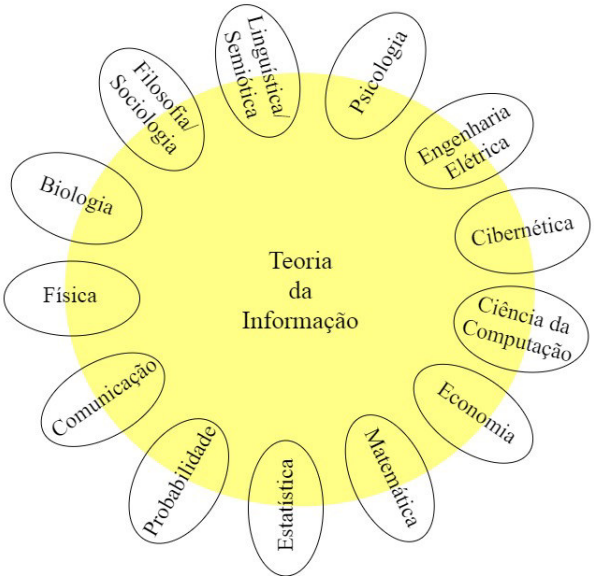
Introdução

Os passos iniciais para a Teoria da Informação ocorreram no ano de 1948, quando o engenheiro e matemático Claude Elwood Shannon (1916-2001) apresentou ao mundo seu importante trabalho intitulado "*A Mathematical Theory of Communication*", conhecido como o "artigo de Shannon" (GALLAGER, 2001). Neste artigo, Shannon apresentou os limites fundamentais na compressão e transmissão confiável de dados em canais ruidosos (MOSER; CHEN, 2012). No início a comunidade ficou surpresa com o artigo, talvez pelo impacto científico, talvez pela falta de compreensão do assunto. Mas com o tempo ele foi aceito e tornou-se a base científica das comunicações, transformado em livro no ano seguinte. Tal credibilidade atribui a Shannon o título de precursor da Teoria da Informação e tornou apropriado que o IEEE chamasse a Teoria da Informação como "Teoria de Shannon" (MASSEY, 1984).

Inicialmente, a Teoria Matemática da Comunicação era entendida como apenas um subconjunto de uma teoria geral de comunicação. Entretanto, era muito mais que isso (COVER; THOMAS, 2012). Com o decorrer do tempo, consolidou-se como um ramo da Teoria da Probabilidade (MOSER; CHEN, 2012) e expandiu-se a outros campos de pesquisa, levando contribuições ou simplesmente mantendo uma relação de proximidade. Há exemplos na física (mecânica estatística), matemática (teoria da probabilidade), engenharia elétrica (teoria da comunicação) e ciência da computação (complexidade algorítmica) (COVER; THOMAS, 2012). Decorrente desta consolidação e expansão, a Teoria Matemática da Comunicação de Shannon identifica-se com um novo campo de pesquisa, atualmente conhecido como Teoria da Informação. A Figura 1.1 ilustra algumas aplicações dessa teoria em outras áreas do conhecimento.

Dois conceitos importantes e revolucionários encontram-se na Teoria Matemática da Comunicação: Teorema de Codificação de Canal e o Teorema de Codificação de Fonte. O Teorema da Codificação de Canal descreve uma maneira de reduzir a taxa de erro em uma transmissão por um canal ruidoso (GAPPMAIR, 1999). Para tornar essa transmissão confiável, a proposta do teorema é que a quantidade de informação a ser transmitida seja sempre menor que a capacidade do canal. Para isso, a determinação da capacidade do canal é muito significativa, pois influencia na escolha da codificação adequada (PIERCE, 1973). No segundo teorema, o Teorema da Codificação de Fonte, é demonstrada a possibilidade de calcular o número de *bits* necessários para descrever um dado de forma única para qualquer fonte de dados (GAPPMAIR,

Figura 1.1 - Relação da Teoria da Informação com outras áreas



Fonte: Adaptada de Cover e Thomas (2012, p. 2).

1999), bem como a ideia de que para transmitir uma quantidade de informação maior, não é necessário aumentar a capacidade do canal, mas sim melhorar a codificação da fonte (MASSEY, 1984). Desta forma, economiza-se tempo e energia (MOSER; CHEN, 2012).

Assim, oriunda destes teoremas destaca-se uma poderosa ferramenta da Teoria da Informação, seu "coração": a codificação (GUIZZO, 2003). A teoria da codificação preocupa-se com métodos de transmissão ou armazenamento de dados de forma eficiente e confiável, composta por técnicas de compressão e controle de erros. Tais técnicas representam a aplicação prática da Teoria da Informação e são constantemente utilizadas em redes de comunicações para reduzir a carga e tornar a transmissão de dados mais robusta aos distúrbios do canal de transmissão (MOSER; CHEN, 2012).

Através da codificação é possível determinar os limites em um processo de comunicação entre máquinas, visando assim à otimização dos recursos. Atualmente, muitas das técnicas de codificação utilizadas tanto para compressão de dados quanto para detecção e correção de erros derivam dos teoremas trazidos da teoria de Shannon (COVER; THOMAS, 2012; MOSER; CHEN, 2012). Apesar de ter predito a codificação, Shannon não forneceu um **algoritmo eficaz** para compressão e não ofereceu algoritmo algum para detecção e correção de erros (GAPPMAIR, 1999).

No entanto, foi a partir da Teoria Matemática da Comunicação de Shannon que surgiram os algoritmos mais comuns utilizados para compressão de dados: o algoritmo de Fano (1949) e o algoritmo de Huffman (1952). Também surgiu o algoritmo de Hamming (1950) para detecção e correção de erros. Desta forma, as leis fundamentais de compressão e transmissão de dados auxiliam no desenvolvimento da Teoria da Informação. Assim, a Teoria Matemática da Comunicação de Shannon pode ser considerada seminal para a abertura de uma nova era: A Era da Informação (VERDU, 1998). A expressão *Annus Mirabilis* (Ano Admirável) é atribuída na história da ciência aos anos nos quais foram realizadas importantes descobertas que afetaram profundamente a humanidade. Dentre eles, o ano de 1665 com Isaac Newton e 1905 com Albert Einstein. Nebeker (1998b) indica que o ano de 1948 pode ser considerado *Annus Mirabilis* graças ao trabalho de Shannon e de outros grandes cientistas que contribuíram para o surgimento e consolidação da Teoria da Informação.

As mudanças decorrentes da Teoria Matemática da Comunicação nas telecomunicações podem ser visualizadas através de uma simples pesquisa na Internet com o termo *"New York Telephone Lines"*. As imagens e documentários exibidos mostram a situação precária do cabeamento telefônico da cidade de *New York* no século XIX, ou seja, a era digital acarretou melhorias nas comunicações. Um segundo exemplo é a busca por fotos de cabos transoceânicos que interligam os sistemas de comunicações dos continentes. Vale lembrar que as comunicações via cabos submarinos eram (e ainda são) muito custosas. Uma significativa mudança ocorreu. Não basta apenas potência e largura de banda para a transmissão de um sinal. Shannon mostrou que é possível atingir o máximo da capacidade de um canal a partir de códigos apropriados. Desta forma, a codificação, compressão e correção de erros são fundamentais para a utilização máxima, próxima aos limites dos canais de transmissão.

As imagens (obtidas na Internet) exibirão o número excessivo de linhas telefônicas e de telégrafos (canais). Até então, a técnica era orientada por conceitos de Armstrong em 1936, de que para diminuir o ruído seria necessário aumentar a banda (PIERCE, 1973). Através do código, Shannon propôs formas de otimização, para utilizar-se o máximo da capacidade da banda já existente, demonstrando que a otimização não deve focalizar-se no aumento da banda e potência, mas sim na utilização de códigos adequados, a partir de um tratamento das informações. Fato este que também possibilita, aliado à codificação para detecção e correção de erros, romper as barreiras que limitavam as transmissões a longa distância, como citado no segundo exemplo (de cabos transoceânicos intercontinentais).

Um terceiro exemplo relaciona-se exatamente às ferramentas de busca na Internet utilizadas para a obtenção dos dois exemplos anteriores. Os conceitos introduzidos por Shannon foram utilizados nestas ferramentas de busca, pois as imagens trazidas no momento da busca na Internet certamente são do formato JPG ou PNG que usam compressão de dados feita por algoritmos baseados na taxa de entropia de Shannon, com eficiência e sem perda relevante de informação (LESNE, 2011). É possível que estejam hospedadas em servidores de outros continentes, e que, com um *click* de *mouse* atravessem oceanos, superando os diversos tipos de ruído, graças a algoritmos de correção de erros, e chegando até a tela de um computador.

Por isso, embora a maioria das pessoas possam não perceber a existência da Teoria da Informação e suas aplicações, os impactos trazidos por ela tornaram possíveis, viáveis e populares a Internet, discos compactos (CD), telefones celulares, *pen drives*, etc. Um grande destaque é o CD, considerado ícone da presença e popularização da Teoria da Informação nos anos 90, pois os dados já eram gravados em formato digital e utilizavam algoritmos tanto para a compressão como para detecção e correção de erros. Inclusive, o mesmo algorítimo utilizado para detecção e correção de erros em CDs foi utilizado na exploração espacial, permitindo às espaçonaves se comunicarem com eficiência mesmo em distâncias longínquas da Terra (MOSER; CHEN, 2012).

A Teoria da Informação também faz-se presente em *modems*, discos rígidos, *chips* de memória, esquemas de criptografia, comunicação óptica, televisão de alta definição, entre outros. Desta forma, é claramente perceptível que a Teoria da Informação não é um assunto antigo, da década de 40, mas sim da contemporaneidade, pois está invisivelmente integrada nas TICs (Tecnologias de Informação e Comunicação) atuais, trazendo mais facilidade e conforto para a vida de seus usuários (GUIZZO, 2003).

Assim, as pessoas estão interagindo com a da Teoria da Informação mesmo que indiretamente. Um exemplo prático talvez seja o meio pelo qual este trabalho de pesquisa possa ser lido. Para ser construído através de texto e imagens, transitar em uma rede de computadores e chegar até os leitores de maneira eletrônica ou impressa, certamente necessitou de conceitos matemáticos trazidos pela Teoria da Informação, como por exemplo: a compressão de dados, a verificação/correção de erros e principalmente a transmissão de sinais. Além deste, há diversos outros exemplos da interação entre as pessoas e a Teoria da Informação sem mesmo conhecê-la. Por exemplo, há "Teoria da Informação" numa conexão pelo *smartphone* à Internet a fim de se comunicar com outras pessoas. Há "Teoria da Informação" ao usar o *smartphone* para tirar uma

simples fotografia ou ouvir uma música. E há certamente algoritmos que realizam a compressão de dados e a detecção/correção de erros nesses dispositivos.

1.1 Motivação

A Teoria da Informação é uma teoria estatística de comunicação com fortes impactos nas tecnologias atuais. Entender os conceitos chaves dessa teoria é importante não só para a ciência como para motivar estudantes interessados em se aprofundar nessa área. A fim de divulgar melhor a Teoria da Informação e a matemática como ciência de base, acredita-se ser interessante e importante neste trabalho um resgate histórico sobre Teoria da Informação.

Considerando-se que a Teoria da Informação indica "o que pode ser feito" e a Teoria da Codificação indica "como fazer" é de se esperar que esses assuntos não sejam tão populares como os produtos tecnológicos deles advindos. Onde há transmissão ou armazenamento de dados, há Teoria da Informação e consequentemente há uma matemática subjacente apropriada aos desenvolvimentos tecnológicos nessa área. De acordo com (GUIZZO, 2003), grande parte das ideias utilizadas hoje nas tecnologias de ponta resultam de leis matemáticas que permitiram que se atingissem os limites de sistemas projetados para transmitir e manipular informação.

Desde que a Teoria da Informação surgiu na década de 40, é notória a evolução da tecnologia. As inovações tecnológicas que por um lado advém de ferramentas matemáticas já desenvolvidas, por outro, propiciam também o surgimento de novos desenvolvimentos matemáticos. Ou seja, é bem possível que problemas abertos em Teoria da Informação, como por exemplo, *Broadcast Channel*, só possam ser resolvidos a partir de novos desenvolvimentos matemáticos.

A Teoria da Informação também é fundamentada em dois conceitos: informação e comunicação. Ambos os conceitos são de uma abrangência e plasticidade muito grande, o que faz com que essa teoria afete, contribua e interaja com diversas áreas, tal como abordado por Cover e Thomas (2012). Esta interação produz interrogações e curiosidades que tornam interessante o estudo de alguns aspectos da Teoria da Informação com fins de esclarecimento. Considerando que a Teoria da Informação elevou o nível da ciência (PINEDA, 2006), subentende-se que o esclarecimento de trechos confusos presentes na literatura pode contribuir para esta elevação, possibilitando uma ciência com mais qualidade. De acordo com Toeplitz (2008), a busca pela gênese de um conceito facilita sua compreensão e possui fins pedagógicos.

É claro que as palavras “comunicação” e “informação” dão margem a inúmeras leituras. Para evitar esse patamar de discussão Shannon deixou claro que sua teoria remete-se aos aspectos sintáticos. Mesmo assim há confusões decorrentes desses termos e de outras palavras utilizadas por Shannon. Por exemplo, o termo “Teoria da Informação” é apontado em alguns textos como sendo de autoria de Shannon e em outros não são abordados (ELLERSICK, 1984). Um outro exemplo refere-se à palavra *entropia* utilizada por Shannon em seu artigo original de 1948. Para que os teoremas de codificação propostos pelo artigo de Shannon se apliquem, é necessário que se desenvolva um critério de medida para a quantidade de informação, escolha e incerteza. Shannon batizou essa medida com o nome de entropia (GAPPMAIR, 1999). A palavra entropia já existia antes da Teoria de Shannon, usada há mais de 80 anos na segunda lei da termodinâmica da física, introduzida por Clausius para denotar a perda de calor que não pode ser convertida em trabalho. O porque da escolha desse nome por Shannon não deixa de ser um assunto de interesse para pesquisa haja vista que há “similaridades” entre a entropia de Shannon e a de Clausius.

Na literatura científica clássica o conceito de entropia é definido de diferentes formas tais como taxa, quantidade e também como unidade de medida, não existindo uma definição padronizada, devido à sua complexidade. Este trabalho adotará a definição de Shannon de entropia como medida de informação, escolha e incerteza (SHANNON, 1948) que é também adotada por alguns dos referenciais teóricos de Teoria da Informação como Reza (1961) e Cover e Thomas (2012).

Mesmo quando o conceito de entropia é limitado à Teoria da Informação, onde a informação é associada à incerteza, existem algumas confusões referentes aos idealizadores de tal conceito. Shannon apresenta sua fórmula, porém Reza (1961), Campbell (1982), Mandrekar e Masani (1997) e Seising (2009) reconhecem a importante contribuição do prof. Dr. Norbert Wiener (1894-1964) na concepção deste conceito na Teoria da Informação.

No mesmo ano em que Shannon publica a Teoria Matemática da Comunicação, Wiener torna-se idealizador da Cibernética. Em suas obras aponta que uma das funções da informação é reduzir a incerteza, definindo o termo *negentropia*, o inverso da entropia. Wiener aborda este conceito de informação *versus* incerteza aplicado a problemas mais gerais de comunicação e controle, abordando desde a comunicação animal até mais especificamente a humana, inclusive aspectos biológicos (WIENER, 1948; WIENER, 1961; WIENER, 1973).

Como mencionado anteriormente, a expansão do cenário da Teoria da Informação para outras áreas do conhecimento tornou comum encontrar muitos trabalhos científicos que referem-se, ou utilizam, os conceitos introduzidos por Shannon (1948). Neste ponto, haja vista a plasticidade e a forte proximidade com outras áreas, e do fato da Teoria da Informação ter 70 anos, ao consultar tais trabalhos, evidencia-se algumas confusões e equívocos acarretando curiosidades que envolvem o conceito de entropia e Teoria da Informação. Além disso, a forte semelhança da Teoria da Informação com processos gerais de comunicação e a utilização da palavra "informação" conduz evidentemente a leituras diversas, sejam elas científicas, sejam elas advindas de contextos históricos. Por exemplo:

1. Conceitos de entropia abordados na Física e na Teoria da Informação. Relação exposta em Ben-Naim (2008) e Ben-Naim (2017) entre a entropia utilizada por Clausius (1864) e a entropia utilizada por Shannon (1948). Com aparência semelhante se comparada ao desenvolvimento da ideia de Clausius por Boltzmann. Há semelhanças e diferenças também apontadas por outros autores como Brillouin (1950), Jaynes (1988), Tribus e McIrvine (1971) e Wehrl (1978).

2. A origem do termo "Teoria da Informação". Apesar do artigo de Shannon de 1948 ter sido o ponto de partida da era da informação e ter sido também apontado como "Carta Magna" da Teoria da Informação (VERDU, 1998), em seu trabalho Shannon não deu ênfase ao termo. O termo ocorre apenas uma vez na página 11. Há, no entanto, a utilização deste termo por Goldman (1948) em um artigo publicado antes do artigo de Shannon. Também, em 1945 é encontrada a utilização deste termo em um relatório técnico de Shannon sobre criptografia. O próprio Shannon em uma entrevista não afirma essa autoria (ELLERSICK, 1984; TRIBUS; MCIRVINE, 1971).

3. Com relação à fórmula da capacidade de canal de Shannon e a regra de Hartley há uma interessante discussão exposta por Rioul e Magossi (2014).

4. É clássico que Shannon é o "inventor" da Teoria da Informação, mas há na literatura menções de "inventores" da Teoria da Informação à Shannon-Weaver, por exemplo em: Albino, Garavelli e Gorgoglione (2004); Al-Fedaghi (2012); Beecher (1989); Cattadori, Haydon e Hudson (2005); Mouillot e Lepretre (1999); Mouton *et al.* (2008); Pak e Paroubek (2010); Shpak e Churchill (2000); Templet (1999); Wang e Li (2017).

5. Também há diversos trabalhos que apontam a criação da Teoria da Informação para Shannon-Wiener. Neste caso, o próprio Wiener afirma que a Teoria da Informação "*[...] as developed by Claude E. Shannon and myself.*" (WIENER, 1956, p. 48). Em uma nota de rodapé na página 34 de seu principal artigo, Shannon cita a importante filosofia e as teorias básicas de Wiener que apontam a comunicação como problema estatístico (SHANNON, 1948). Wiener também cita Shannon

   > *This is precisely the result which the author and Shannon have already obtained for the rate of transmission of information in this case*" (WIENER, 1961, p. 87).

   A relação Shannon-Wiener também é abordada no livro "*Dark Hero of the Information Age: In Search of Norbert Wiener, The Father of Cybernetics*" destacando a importância de Wiener na Teoria da Informação (CONWAY; SIEGELMAN, 2006).

1.2 Objetivos

Objetiva-se realizar uma releitura do conceito de entropia em Teoria da informação de forma que possibilite indicar um "caminho", fundamentado em asserções históricas contidas em livros e artigos, que elimine algumas contradições linguísticas e facilite o "andar científico" em áreas da Teoria da Informação dependentes desse conceito. Esse caminho torna-se possível graças ao acesso às informações e às facilidades tecnológicas disponíveis atualmente tais como Internet, bases de dados, artigos etc. Desta forma, os objetivos específicos desta pesquisa são:

1. Realizar um resgate histórico, também com fins pedagógicos, visando a reduzir "incertezas" sobre a ocorrência do conceito de entropia na Teoria da Informação.

2. Indicar um caminho histórico onde haja a descrição da gênese dos conceitos essenciais à Teoria da Informação, importante para a compreensão de problemas atuais.

3. Identificar um problema aberto em Teoria da Informação que esteja alicerçado no conceito de entropia.

1.3 Propostas na dissertação

Na literatura científica, há uma fundamentação matemática sólida referente à Teoria da Informação. Pode-se citar Reza (1961), Cover e Thomas (2012), Kinchin (1957) entre

outros[1]. Há também uma diversidade de trabalhos que resgatam conceitos históricos de Teoria da Informação, bem como indicam um refinamento na definição de conceitos, como exemplos: Brillouin (1950), Cherry (1951), Pierce (1968), Massey (1984), Aspray (1985), Nebeker (1998a), Verdu (1998), Gappmair (1999), Gallager (2001), Lundheim (2002), Golomb *et al.* (2002), Guizzo (2003), Pineda (2006), Ben-Naim (2010) e Moser e Chen (2012). Além destes, há ainda outros autores que abordam mais especificamente o conceito de entropia. No entanto, observou-se que, apesar de alguns destes trabalhos apontarem para tecnologias atuais, nos quais estão presentes os conceitos trazidos pelo artigo de Shannon (1948), e clarearem algumas confusões pontuais existentes em Teoria da Informação, não focalizam exclusivamente nestas confusões, deixando assim uma lacuna que permite a ocorrência de leituras múltiplas associadas aos conceitos de Teoria da Informação.

No capítulo 2, a discussão contextualizará o ambiente em que ocorrem as confusões apontadas como objeto de investigação, por isso fará uma caracterização da Teoria da Informação, abordando os importantes teoremas trazidos de Shannon, a importância da entropia para a aplicação destes teoremas, bem como apresentando um resgate histórico da origem da palavra entropia e da simbologia utilizada para representá-la. Também há uma abordagem sobre a utilização da entropia além da física e da Teoria da Informação. Desta forma, busca-se exibir uma discussão com fins de esclarecimento dos problemas que envolvem o conceito de entropia. Neste capítulo, há também uma abordagem sobre a relação da Teoria da Informação com outras áreas da ciência, bem como uma breve biografia de Claude Elwood Shannon.

No capítulo 3 expõe-se, de modo breve, um pouco da Teoria de Códigos e sua relação com Teoria da Informação. Neste capítulo expõe-se ainda um interessante problema presente em muitos setores do mercado tecnológico, qual seja, o problema do canal de difusão (*Broadcast Channel*). A escolha desse problema se deve ao fato de que, nele, o conceito de entropia é apresentado para ser utilizado de forma múltipla, além daquilo que foi apresentado por Shannon.

Não há como negar a genialidade de Shannon ao escolher um modelo matemático para comunicação. Essa escolha possibilitou o surgimento das atuais Tecnologias de Informações e Comunicações fortemente fundamentadas em seu modelo matemático.

[1] Ver por exemplo o livro de Verdú, McLaughlin e Society (2000)

1.4 Contribuições da Dissertação

Com o propósito de compartilhamento e divulgação acadêmica/científica, alguns resultados preliminares obtidos durante a pesquisa e abordados nesta dissertação foram publicadas:

1. Apresentação de pôster no X Workshop da pós-graduação da FT/Unicamp, bem como publicação *online* nos anais deste evento: PAVIOTTI, José Renato; MAGOSSI, José Carlos. *Considerações sobre o conceito de entropia na Teoria da Informação*. In: WORKSHOP DA PÓS-GRADUAÇÃO DA FACULDADE DE TECNOLOGIA, 10, 2018, Limeira/SP. Anais 2018, 2018, p. 24. Disponível em: <https://wordpress.ft.unicamp.br/workshoppos/edicao-2018/>. Acesso em: 30 out. 2018 .

2. MAGOSSI, José Carlos; PAVIOTTI, José Renato. Incerteza em Entropia. *Revista Brasileira de História da Ciência*. **Submetido em 12/2018** .

Capítulo 2

Resgate histórico dos conceitos em Teoria da Informação

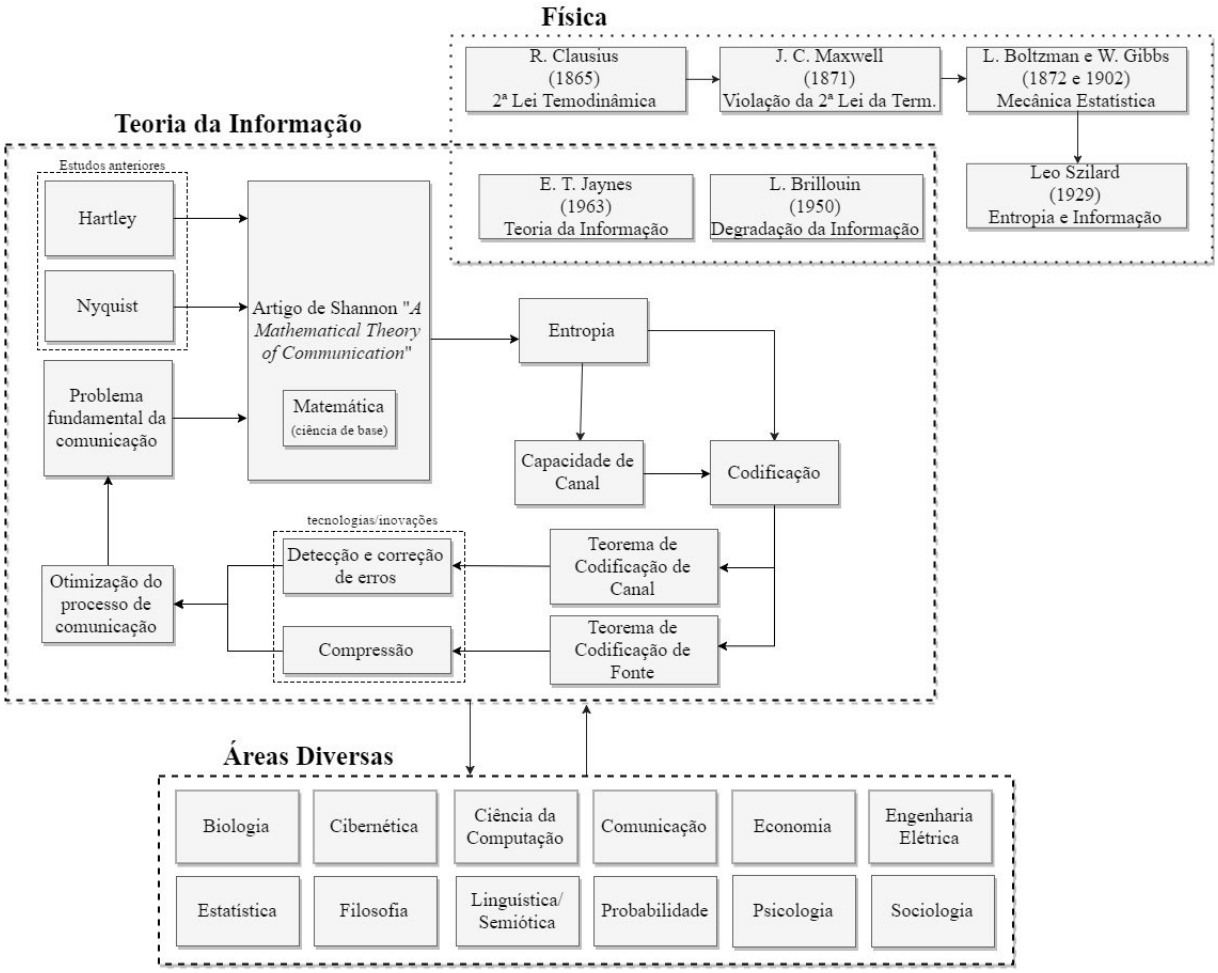
Como já descrito no capítulo anterior, um dos objetivos deste trabalho é fazer um breve resgate histórico dos principais conceitos em Teoria da Informação, com o intuito de indicar sua gênese histórica e possibilitar o esclarecimento de possíveis leituras adversas. É claro que a obra de Shannon (1948), apesar de ser pouco conhecida nos meios não acadêmicos, desperta interesse em sua exploração para o desenvolvimento de tecnologia e inovação. Desta forma, a discussão neste capítulo inicia-se com a contextualização da Teoria da Informação de modo a expor a importância do modelo matemático de Shannon para o desenvolvimento de tecnologias tais como *smartphone*, TV digital, Internet, CD, DVD, *pendrive* etc. Por exemplo, a estatística, fortemente desenvolvida no início do século XX, é fundamental para os trabalhos de Shannon. Dessa forma questiona-se se seria possível uma teoria matemática da comunicação em períodos anteriores ao século XX.

Este capítulo começa com uma breve biografia de Shannon. A seguir, aborda a Teoria da Informação sob a ótica da Teoria Matemática da Comunicação de Shannon, a matemática presente nos importantes teoremas de sua teoria, ressaltando o conceito de entropia, apontado como marco zero ou alicerce para a construção da Teoria da Informação. Na conclusão deste capítulo, há uma seção que expõe de uma maneira ampla a relação da Teoria da Informação com outras áreas do conhecimento. A Figura 2.1 representa, através de um mapa mental[1], a proposta de discussão deste capítulo. É importante destacar que a entropia também é o elo de ligação entre Teoria da Informação e outros domínios da ciência, por isso serão apresentadas e discutidas algumas destas relações específicas.

Também é importante apontar as delimitações da proposta de contextualização da Teoria da Informação deste capítulo: elas estão direcionadas às contribuições de Shannon ou às associadas diretamente às dele. Outras contribuições, bem como considerações relacionadas a elas, são abordadas apenas no aspecto histórico desta dissertação. Considerando a complexidade e plasticidade dos conceitos de Teoria da Informação já apontados no capítulo anterior, em nenhum momento busca-se completude na exposição.

[1] Diagrama criado por Tony Buzan para a representação gráfica de ideias, em que estas possuam conexões entre si e almejem facilitar a compreensão de um determinado tópico.

Figura 2.1 - Mapa mental das discussões do capítulo 2



Fonte: Produção do próprio autor.

2.1 Breve biografia de Claude Elwood Shannon

Claude Elwood Shannon nasceu em 30 de abril de 1916 em uma pequena cidade do estado de Michigan chamada Petoskey (GAPPMAIR, 1999; GALLAGER, 2001; MOSER; CHEN, 2012). Na infância, estudou em uma escola pública na qual sua mãe fora professora de línguas e diretora (PINEDA, 2006). Ainda na infância produziu inventos interessantes e curiosos como, por exemplo, um barco controlado por rádio e um telégrafo usando arames farpados de uma cerca entre sua casa e a casa de um vizinho (PINEDA, 2006).

Na vida acadêmica graduou-se em matemática e engenharia elétrica pela Universidade de Michigan (GALLAGER, 2001). Concluiu seu mestrado em 1937, orientado por Vannevar Bush (GALLAGER, 2001; MOSER; CHEN, 2012). Tratava-se de uma pequisa sobre a utilização da álgebra booleana para comutação de circuitos com relés, com título de "*A Symbolic*

Analysis of Relay and Switching Circuits" (GALLAGER, 2001; MIT, 2001). Com este trabalho ele mostrou como os circuitos elétricos podem ser utilizados para realizar operações lógicas e matemáticas. Tal trabalho tornou-se fundamental para a computação digital e para a teoria dos circuitos digitais, ferramenta essencial para a indústria microeletrônica do século XIX, que possibilitou a tecnologia da informação e comunicação (TIC) do século XXI (MOSER; CHEN, 2012). No doutorado, Shannon optou pela área da genética. Trabalhou em uma tese intitulada de "An Algebra for Theorical Genetics" (GALLAGER, 2001). Sua tese foi concluída em menos de um ano, porém decidiu não permanecer neste campo de pesquisa (GUIZZO, 2003).

Embora tivesse focalizado em álgebra tanto no mestrado quanto no doutorado, Shannon interessou-se muito por telefonia e tinha uma grande preocupação com os sistemas de comunicação (GALLAGER, 2001). Durante a segunda guerra mundial, em 1941, juntou-se à *Bell Laboratories* (GUIZZO, 2003). Logo após o término da segunda guerra mundial, ainda durante o período que esteve afiliado à *Bell Laboratories*, precisamente no ano de 1948, Shannon publica seu artigo mais importante: "*A Mathematical Theory of Communication*" (GAPPMAIR, 1999).

No ano de 1956 deixou os Laboratórios Bell e foi convidado a ingressar no MIT como professor visitante e no ano seguinte tornou-se membro pleno, onde permaneceu até sua aposentadoria em 1978 (GALLAGER, 2001; GAPPMAIR, 1999). O Prof. Dr. Claude Elwood Shannon recebeu títulos *honoris causa* pela Universidades de Yale, Michigan, Princeton, Edimburgo, Pittsburgh, Northwestern, Oxford, East Anglia, Carnegie-Mellon, Tufts e da Pensilvânia. (MIT, 2001).

Outros trabalhos curiosos e interessantes foram realizados por Shannon os quais tornam possível entender um pouco sua genialidade e criatividade, bem como seu conhecido senso de humor e humildade. Dentre seus trabalhos, podem ser citados (GALLAGER, 2001; GAPPMAIR, 1999; GUIZZO, 2003):

- A programação de computador para jogar xadrez (GUIZZO, 2003).
- Theseus: um rato mecânico capaz de atravessar um labirinto (MIT, 2001; GUIZZO, 2003).
- Penny Matching: se caracterizava como um algoritmo que procurava o padrão de jogo do adversário (MIT, 2001).

Tais inventos não deixam esconder o gosto e a boa relação de Shannon com o xadrez. Shannon também inventou a máquina inútil, "*Useless Machine*"[2]. Este invento trata-se de uma caixa fechada que, ao se acionar um interruptor externo à caixa, a caixa abre-se e de dentro dela sai um dispositivo mecânico (parecido com um braço) que aciona o interruptor para novamente fechá-la. Shannon também tentou criar uma máquina para resolver o cubo de Rubik (MIT, 2001; GUIZZO, 2003). Ele não conseguiu inventar a máquina, mas criou uma música sobre o cubo. Atualmente existem robôs, até mesmo construídos em plataformas pedagógicas como Arduíno[3] que fazem isso. Outra tentativa curiosa foi a de criar uma teoria matemática para o malabarismo. Submeteu o trabalho à revista *Scientific American*, a qual solicitou revisões à Shannon, o que fez com ele desistisse da publicação (GUIZZO, 2003).

Durante a segunda guerra mundial Shannon trabalhou com criptografia. A relação da criptografia com a guerra advém provavelmente de motivações militares. No entanto, quando questionado pelo Dr. Robert Price sobre esta motivação, Shannon negou e disse que este assunto sempre o atraiu e desde menino gostava de resolver criptogramas. Foi durante estes trabalhos com criptografia que ele fundamentou parte da sua teoria matemática para comunicação (ELLERSICK, 1984).

Shannon casou-se em 1949 com Mary Betty Moore, matemática da *Bell Labs* que trabalhava no grupo de John Pierce. Após se casarem, foram morar em Nova York e depois em Nova Jersey. Shannon tornou-se professor emérito do MIT em 1978 e passou seu tempo em casa, a maior parte brincando com seus filhos (GUIZZO, 2003).

Na década de 80, Shannon foi diagnosticado com Alzheimer (GALLAGER, 2001). A doença trouxe grandes dificuldades nos anos seguintes. Na década de 90 já não conseguia encontrar o caminho de volta para casa. Com o passar do tempo ele não reconhecia mais seus próprios escritos e amigos (GUIZZO, 2003). Considerada sua morte como uma grande perda para a comunidade científica, Shannon lutou contra o Alzheimer até o dia 24 de fevereiro de 2001, quando faleceu próximo de completar 85 anos de idade (GALLAGER, 2001).

[2]Disponível em vídeos na Internet. Por exemplo em: <https://www.youtube.com/watch?v=G5rJJgt_5mg> e <https://www.youtube.com/watch?v=eeC2aXXPSQw>.

[3]Plataforma livre que permite a prototipagem eletrônica de hardware. É composta por uma uma placa que possui um microcontrolador com suporte de entrada e saída que é controlado por uma linguagem de programação.

2.2 A Teoria Matemática da Comunicação de Shannon

Logo que publicada, a Teoria Matemática da Comunicação de (SHANNON, 1948) causou grande espanto e interesse entre os engenheiros elétricos e interessados no campo de comunicação, pois o artigo apresentou uma teoria que propunha um modelo matemático para a comunicação, fato este que não a relacionava exclusivamente a uma tecnologia particular. Visto como um artigo completamente fora dos padrões atuais, o artigo de Shannon intitulado de "Uma Teoria Matemática para Comunicação" apresenta 23 teoremas, 7 apêndices com provas matemáticas ao longo das 77 páginas. Talvez devido à sua extensão, tenha sido publicado em duas partes nas edições de julho e outubro de 1948 da revista técnica dos Laboratórios Bell (GUIZZO, 2003).

Como revelado pelo próprio Shannon (1956) em "*The bandwagon*", a proposta inicial do trabalho "*A Mathematical Theory of Communication*" (SHANNON, 1948), não foi almejada de maneira tão audaciosa a fim de afetar diversas áreas do conhecimento e criar uma nova forma de pensar a informação e a comunicação. A proposta inicial do trabalho de Shannon surgiu de maneira modesta, como uma forma matemática de resolver o problema fundamental da comunicação, qual seja, o de "[...] reproduzir em um ponto exatamente ou aproximadamente uma mensagem selecionada em outro ponto" [4] (SHANNON, 1948, p. 1, tradução nossa). No entanto, o artigo de Shannon iniciou uma nova era, na qual a informação é vista por ângulos diferentes, dando início à Teoria da Informação (VERDU, 1998).

De acordo com Guizzo (2003), diferentemente de outros casos na história da ciência cujos momentos de *insights* ficaram eternizados como, por exemplo, o momento "eureka" de Arquimedes, a Teoria da Informação não tem nenhum ícone simbólico em sua origem, mas tem um início muito claro, fundado por Shannon através da publicação de seu artigo de 1948. Tal trabalho é considerado uma obra-prima, pois

> "Antes de Shannon, os engenheiros não tinham respostas claras a essas questões. Naquela época, um zoológico selvagem de tecnologias estava em operação, cada uma com uma vida própria - telefone, telégrafo, rádio, televisão, radar e vários outros sistemas desenvolvidos durante a guerra" (GUIZZO, 2003, p. 8, tradução nossa)[5].

[4] *"The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point."*

[5] *"Before Shannon, engineers had no clear answers to these questions. At that time, a wild zoo of technologies was in operation, each with a life of its own - telephone, telegraph, radio, television, radar, and a number of other systems developed during the war."*

Por isso, a Teoria Matemática da Comunicação de Shannon foi considerada inovadora, sintetizando os conceitos de comunicação e informação abstraídos a partir de conceitos matemáticos universais, aplicados a qualquer tipo de canal, sinal ou dado que se desejasse transmitir. Portanto, apesar do termo Teoria da Informação ter se difundido alguns anos depois, a partir da ramificação da Teoria Matemática da Comunicação a outros campos de pesquisa, a Teoria da Informação e a Teoria de Shannon tornaram-se indissociáveis.

A Teoria da Informação consegue extrapolar mesmo quando vista a partir de uma concepção simplória e limitada, como a de um subconjunto da Teoria da Comunicação. Exemplo desta extrapolação é a forma como afeta o próprio processo "natural" de comunicação. Tal processo pode ser representado através de um diagrama de blocos composto por 3 partes essenciais: Transmissor (ou fonte), Canal e Receptor (Figura 2.2). Após a nova visão trazida por Shannon, esse modelo simples e limitado é expandido, visto de uma maneira mais ampla, acrescentando-se novos elementos, como observa-se na Figura 2.3 (REZA, 1961).

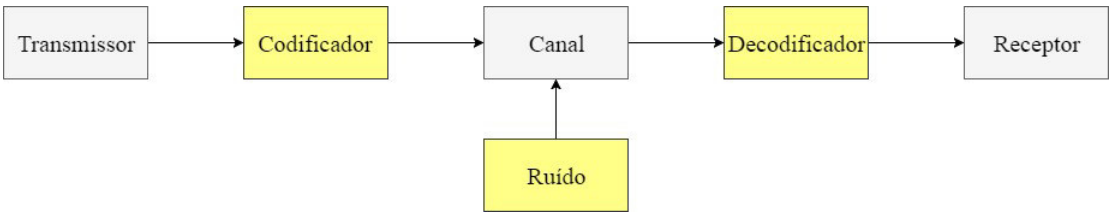
Neste modelo, o primeiro componente é o transmissor ou fonte de dados. Trata-se do local em que a mensagem é originada e codificada em uma forma adequada para transmissão. O segundo componente é o canal. Através dele, a mensagem codificada é enviada ao destinatário. Durante a transmissão, a mensagem pode ser afetada por erros (COVER; THOMAS, 2012). Tais erros são caracterizados por distorções aleatórias na mensagem chamados tecnicamente de ruídos (BROOKES, 1956). De acordo com Eco (2001, p. 7), o "ruído é um distúrbio que se insere no canal e pode alterar a estrutura física do sinal".

Figura 2.2 - Diagrama de blocos: Modelo básico de um sistema de comunicação

Fonte: Adaptada de Reza (1961, p. 2).

Na Figura 2.3, podemos observar que além das 3 partes essenciais observadas na Figura 2.2, existem 3 novos componentes apontados por Shannon em seu artigo: o codificador, o decodificador e a presença de ruído no canal. O principal foco era o desafio de superar o problema do ruído, cuja solução representou um dos grandes diferenciais da teoria de Shannon. Outras importantes descobertas semelhantes da época tentavam atenuar o ruído ou até mesmo o desconsideravam, sendo Shannon o pioneiro a provar que é possível fazer uma transmissão sem erros através de um canal ruidoso (LUNDHEIM, 2002).

Figura 2.3 - Modelo de um sistema de comunicação que utiliza a Teoria da Informação



Fonte: Adaptada de Reza (1961, p. 4).

Outra grande ideia trazida pela Teoria Matemática da Comunicação propunha que toda comunicação devesse ocorrer de forma digital, equivalente na geração, transmissão e recepção (MASSEY, 1984), baseada em uma nova unidade de informação: o *bit*, expressão formada pela junção das palavras *binary* e *digit*. Tal expressão foi sugerida a Shannon por J. W. Tukey (SHANNON, 1948).

A informação vista de forma digital ampliou muitos horizontes, muitos deles pelo processamento digital de sinais. Até meados dos anos 60 os sinais eram convertidos em forma de ondas de corrente ou tensão. Hoje em dia, os sinais são convertidos para o formato digital e são mais facilmente compreendidos e manipulados por processadores. É mais eficiente trabalhar com sinais digitais do que com analógicos, pois oferecem vantagens como a flexibilidade na construção dos circuitos, a reconfiguração de um sistema eletrônico e principalmente pelo fato de serem mais baratos (NALON, 2009)

Mesmo com as novas ideias sendo rapidamente aceitas pela comunidade científica, elas demoraram para serem aceitas e colocadas em prática pelos engenheiros de comunicação, pois alguns se fundamentavam no modelo clássico da antiga Teoria da Comunicação, já utilizado por décadas, que baseava-se em técnicas de modulação analógicas e relação de maior banda/menor ruído, por exemplo (MASSEY, 1984). De acordo com Massey (1984), a Teoria da Informação é parecida com a Teoria de Copérnico. Ambas foram inicialmente desacreditadas, porém comprovadas a posteriori. De acordo com Guizzo (2003) isto é comum e característico das teorias revolucionárias, pois forçam os cientistas e pesquisadores a repensar o que consideravam verdadeiro e fundamental.

Logo no ano seguinte ao da publicação da Teoria Matemática da Comunicação, em 1949, fortes críticas foram publicadas em um artigo de revisão de Doob (1949). Joseph L. Doob (1910-2004), era um grande matemático que tornou-se referência na área de processos estocásticos. Suas críticas apontavam para a falta de rigor matemático e ausência de demonstrações

suficientes (GOLOMB *et al.*, 2002) e (GUIZZO, 2003). Realmente havia falhas em teoremas e demonstrações. Na teoria de Shannon, como em toda teoria científica, há desdobramentos não previstos em sua concepção original. A matemática subjacente à Teoria da Informação ainda hoje é estudada e com problemas abertos sendo investigados. No artigo original de Shannon alguns problemas de codificação não foram resolvidos. O prof. Dr. Robert Fano, em suas aulas, desafiava seus alunos a resolver alguns problemas presentes no artigo de Shannon. Desses desafios e discussões, algoritmos para codificação foram elaborados por R. Fano. Como outro exemplo pode-se citar a demonstração do teorema da Capacidade de Canal de Shannon, desenvolvida pelo matemático McMillan (GUIZZO, 2003).

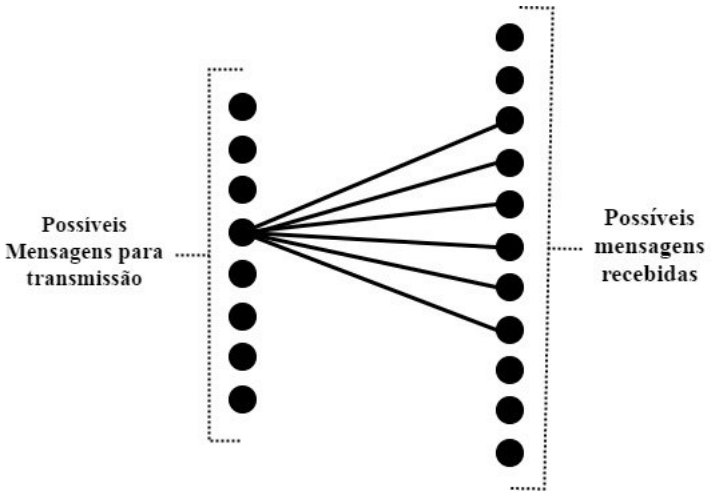
Nos anos posteriores muitos trabalhos foram publicados endossando e provando o sucesso da teoria de Shannon. Dentre eles o trabalho de Woodward e Davies (1952) sugerindo uma forma de construção de aparelhos receptores (radares) utilizando conceitos estatísticos da teoria de Shannon. Já na área de codificação, alguns dos principais desdobramentos da Teoria da Informação, surgem dos trabalhos de Fano (1949), Hamming (1950) e Huffman (1952). Tais trabalhos serão explorados com maior amplitude nas seções seguintes desta dissertação.

Entre os matemáticos, a teoria de Shannon começou a ser aceita e compreendida a partir da publicação do artigo de McMillan (1953). Shannon, enquanto produzia sua teoria, consultou o experiente matemático Brockway McMillan solicitando ajuda com codificação. No entanto, evitando a exposição na íntegra de seu trabalho ao colega, Shannon tentou desenhar um diagrama com pontos e linhas indicando o envio e recebimento de informações. Inicialmente o desenho havia ficado muito confuso, parecendo uma hélice de ventilador, por isso foi chamado de "diagrama ventilador". Na Figura 2.4 há uma ilustração de tal diagrama. Tal confusão e falta de contextualização adequada, fez com que McMillan não entendesse o problema e, em um primeiro momento, não pudesse ajudá-lo. No entanto, logo após a publicação do artigo de Shannon, McMillan entendeu os tais diagramas de ventilador e motivou-se a escrever um artigo[6] (GUIZZO, 2003).

A revolução microeletrônica iniciou-se quase que na mesma época que Shannon criava a Teoria da Informação, pois fora também iniciada na Bell Labs, em 1947, quando os pesquisadores Walter Brattain, John Bardeen e William Shockley inventaram o transistor (GUIZZO, 2003). No entanto, somente a partir na década de 70, iniciou-se a produção em massa de circuitos digitais e computadores (GAPPMAIR, 1999; GUIZZO, 2003). Somente

[6]McMillan (1953)

Figura 2.4 - Diagrama “ventilador” de Shannon



Fonte: Adaptada de Guizzo (2003, p. 42).

com o avanço da eletrônica, e paralelamente dos circuitos digitais, é que as ideias de Shannon puderam ser implementadas. Tal atraso causou medo de que a teoria de Shannon fosse perdida (GALLAGER, 2001). Talvez esse seja o motivo do comentário de Pierce (1973), ao considerar que as ideias do artigo de Shannon caíram como uma bomba, mas com efeito retardado. Por isso, a lentidão na produção de baixo custo de componentes eletrônicos associou-se também a lentidão na implementação prática da teoria de Shannon.

O processamento, transmissão e armazenamento de sinais digitais demorou um certo tempo para ser realmente aceito e utilizado, sendo implementado ao longo de décadas. Um artigo escrito por Nebeker, intitulado de “*Fifty Years of Signal Processing: The IEEE Signal Processing Society and its Technologies 1948-1998*”, descreve cronologicamente e historicamente este processo. Nos tópicos a seguir apresenta-se uma síntese destas implementações nas décadas de 60, 70, 80 e 90.

- Década de 60: O conceito digital iniciou-se com o processamento de dados em 1961, quando a *Texas Instruments* criou o primeiro computador transistorizado, o TI187. Em meados desta década, a descoberta do algoritmo da FFT[7] por Cooley e Tukey (1965) para transformação rápida de sinais do domínio do tempo para o domínio da frequência estimulou os estudos sobre processamento de sinais (NEBEKER, 1998b). Nessa época, final da década de 40 e década de 50, investiu-se exaustivamente em resultados teóricos, haja

[7] *Fast Fourier Transform* (Transformada Rápida de Fourier)

vista serem escassos os recursos para armazenamento e transmissão digital (GUIZZO, 2003).

- Década de 70: As primeiras transmissões digitais começaram logo nos anos iniciais da década de 70, na França e posteriormente nos Estados Unidos. Impulsionado pela FFT, a década de 70 foi marcada pelo surgimento da filtragem de sinais digitais. Associando o PCM[8] com a filtragem adaptativa de sinais por exemplo, a britânica BBC já fazia transmissões digitais de áudio de alta qualidade para seus sistemas de rádio e TV (NEBEKER, 1998b).

- Década de 80: Esta década é marcada pelo surgimento do armazenamento digital através do CD. O processamento de imagens também recebeu uma atenção maior dos engenheiros, resultando em importantes inventos tais como: fax, tomografia computadorizada e ressonância magnética por imagens. Também nos anos 80 surgiu o JPEG, um padrão para digitalização e compressão de imagens que automaticamente despertou o interesse por estudos de reconhecimento de padrões em imagens (NEBEKER, 1998b).

- Década de 90: Impulsionado pelo crescimento da rede mundial de computadores e do serviço WWW que possibilitou a transmissão de multimídia pela Internet, inspirado no JPEG, surgem as 4 versões do MPEG, um padrão para vídeos digitais. A versão 1 apresentava uma sintaxe para a transmissão de vídeo em *broadcast*. A versão 2 trazia uma padronização para TV Digital e DVD e a versão 3 era apenas um aperfeiçoamento da versão 2, padronizando a TV digital de alta definição (HDTV). A versão 4 traz a padronização de vídeo digital para conexões da Internet (NEBEKER, 1998b).

2.2.1 O termo Teoria da Informação e sua relação com o artigo de Shannon

A associação do termo "Teoria da Informação" com o artigo de Shannon (1948) é facilmente observada. Pode ser notada ao realizar buscas em bases de indexação de periódicos conceituados. Ao observar os resultados encontrados é possível notar que antes da publicação do artigo de Shannon, em julho de 1948, é praticamente raro encontrar referências que usem o

[8] *Pulse Code Modulation* (Modulação por Código de Pulsos). Na seção seção 2.3.1 é feita uma abordagem explicativa sobre PCM.

termo Teoria da Informação. As primeiras publicações na literatura científica relacionada que usam este termo surgem ainda em 1948 e se estendem pelos anos seguintes, como exemplos:

- 1948: "*Network Theory Comes of Age*" (DIETZOLD, 1948); "*Cybernetics*" (WIENER, 1948).
- 1949: "*Life, Thermodinamics Cybernetcs*" (BRILLOUIN, 1949); "*Acoustics in Communication*" (BOWN, 1949).
- 1950: "*Thermodynamics and Information Theory*" (BRILLOUIN, 1950);"*The Information Theory Point of View in Speech Communication*" (FANO, 1950); "*The Intelligibility of Amplitude-Dichotomized, Time-Quantized Speech Waves*" (LICKLIDER, 1950); "*Language Engineering*" (MILLER, 1950); "*The Theory of Information*" (REICH, 1950).
- 1951: "*A history of the theory of information*" (CHERRY, 1951); "*The Theory of Information*" (BARNARD, 1951)

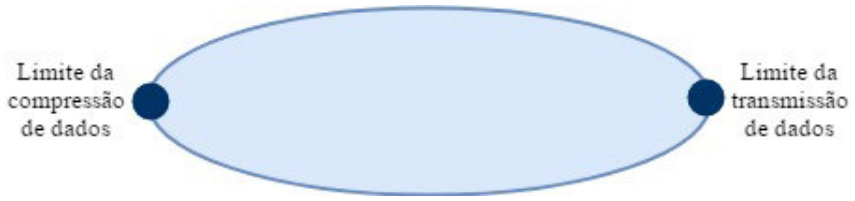
Uma curiosa exceção ao período relacionado acima, é o uso do termo Teoria da Informação dois meses antes do artigo de Shannon de 1948. Trata-se do trabalho de Goldman (1948), endossado por um relatório técnico do MIT datado do ano de 1947 que apontam os trabalhos realizados pelos seus pesquisadores, dentre eles o do Dr. Stanford Goldman, citando o "*general analysis based upon information theory and the mathematical theory of probability is used to investigate the fundamental principles involved in the transmission of signals, through a background of random noise*" (GOLDMAN *et al.*, 1947, p. 46). O artigo de Goldman (1948) cita o termo "Teoria da Informação" logo em seu resumo. No entanto, entende-se que o termo é empregado de maneira genérica, pois seu estudo tem por objetivo investigar princípios fundamentais envolvidos na transmissão de sinais sob ruídos aleatórios, e provar três algoritmos que fazem as relações de probabilidade entre sinal e ruído em transmissões por radar.

No artigo de Shannon de 1948, o termo Teoria da Informação aparece sem muita ênfase na página 11, quando a fórmula da entropia é explicada. Embora Shannon não tenha mencionado o termo Teoria da Informação com muito impacto em seu artigo de 1948, Ellersick (1984), aponta-o como propositor deste termo, fundamentando sua conclusão em um trabalho realizado por Shannon em 1945, intitulado de "*A Mathematical Theory of Criptography*" (SHANNON, 1945). De acordo com Ellersick (1984) e Golomb *et al.* (2002) este trabalho de Shannon em 1945 foi publicado posteriormente, em 1949, como "*Communications Theory of*

Secrecy Information" (SHANNON, 1949). Apesar de Shannon não reconhecer sua autoria na origem do termo Teoria da Informação, como também não considerar sua teoria como a Teoria da Informação, somente durante e após 1948 este termo é amplamente utilizado e associado às descobertas de Shannon. Ao ser questionado sobre esse trabalho de 1945, Shannon reconheceu a importância do mesmo para a construção da Teoria da Informação, relatando inclusive que este trabalho prévio foi realizado em um momento em que ele ainda não se considerava apto a escrever uma teoria matemática da comunicação (ELLERSICK, 1984).

Finalizando esta contextualização, a grande contribuição de Shannon para a Teoria da Informação foi mostrar a existência de uma medida de informação, independente dos meios utilizados para gerá-la. A partir dela, foi possível estabelecer a capacidade de transmissão em um canal e determinar métodos de codificação para otimização deste processo, o que torna facilitada a tarefa de engenharia para projetar melhorias nos canais de comunicação (TRIBUS; MCIRVINE, 1971). Com base neste princípio na Figura 2.5, expõe-se a proposta de otimização do processo de comunicação trazido pela Teoria da Informação, de modo que seja possível a aproximação entre si dos pontos extremos de compressão e transmissão de dados (COVER; THOMAS, 2012). Na comunicação, não é importante apenas saber onde estão as limitações, mas também o quão perto pode-se chegar delas (MOSER; CHEN, 2012).

Figura 2.5 - Ideia de otimização trazida pela Teoria da Informação



Fonte: Adaptada de Cover e Thomas (2012, p. 2).

2.2.2 Shannon-Weaver e o livro *A Teoria Matemática da Comunicação*

Uma recaracterização da Teoria Matemática da Comunicação é feita por C. E. Shannon e W. Weaver (SHANNON; WEAVER, 1964) quando trocam a letra A, artigo indefinido em inglês, do artigo de Shannon, *A Mathematical Theory of Communication* por *The*, artigo definido em inglês, no livro *The Mathematical Theory of Communication*, publicado em sua primeira edição, cronologicamente após os artigos de Shannon (1948) e Weaver (1949).

Nesse livro, a primeira parte é dedicada a uma visão extensiva da Teoria Matemática da Comunicação e suas possíveis implicações práticas. Inicialmente é feita uma abordagem de como uma mente afeta outra e a informação é classificada nos níveis: técnico, semântico e influente. No nível técnico é discutida a aplicação da teoria de Shannon, onde a preocupação é a precisão e bidirecionalidade da informação. O nível semântico preocupa-se com a linguagem pela qual será interpretada a informação no receptor. Por fim, o nível influente destaca o poder da informação de afetar a conduta do receptor. Tais conceitos também foram abordados anteriormente no artigo de Weaver (1949).

Somente na segunda parte deste livro encontra-se a teoria de Shannon (SHANNON, 1948). Uma leitura descuidada pode levar o leitor a confundir a teoria de Shannon com o primeiro capítulo do referido livro. Por esse motivo, provavelmente, é que em muitos textos encontra-se a denominação dos "inventores" da Teoria da Informação como sendo Shannon-Weaver.

No livro *La Théorie Mathématique de la Communication*, recente tradução francesa do clássico Shannon-Weaver, publicada na França pela editora Cassini, o professor Olivier Rioul, no prefácio (páginas 1 a 5), faz uma explicação dessa confusão tão comum e propõe que os capítulos do livro sejam invertidos na tradução francesa de modo que a teoria de Shannon ocorra logo no início e a parte escrita por Warren Weaver fique como apêndice.

Talvez ainda como reflexo desta confusão, há uma outra confusão ainda mais específica, porém muito comum. Trata-se da citação da fórmula da entropia de Shannon (Equação 2.1 adiante) como fórmula de Shannon-Weaver. Portanto, há conceitos da Teoria de Shannon sendo usados por outras áreas com clareza, porém também existem confusões e equívocos que facilmente são encontrados nas publicações científicas. Seguem alguns exemplos:

1. Beecher (1989), Cattadori, Haydon e Hudson (2005) e Pak e Paroubek (2010) : Nestas obras os autores citam o livro de Shannon e Weaver, porém tem todo um cuidado ao apontar para a fórmula da entropia como sendo de Shannon.

2. Mouton *et al.* (2008, p. 125): O autor cita a entropia como sendo de autoria de Shannon-Weaver na página 125: "*The Shannon-Weaver entropy [...]*".

3. Templet (1999): Em vários pontos o autor apresenta o termo "*Shannon-Weaver equation*". Na página 225 apresenta a fórmula da entropia de Shannon como de Shannon-Weaver.

4. Shpak e Churchill (2000): Na seção de resultados, página 235, os autores citam a entropia como sendo de Shannon-Weaver quando discutem resultados de uma figura: "*Shannon–Weaver entropy is shown as a function of rate in Fig. 2 for [...]*". No entanto, na legenda da figura referida, localizada na página 236, a entropia é apresentada como sendo de Shannon: "*Shannon entropy is shown[...]*".

Além da "incerteza" existente sobre a origem do termo "Teoria da Informação", os questionamentos sobre a associação da Teoria de Shannon com a Teoria da Informação e os "inventores" da Teoria da Informação, há uma discussão ainda mais profunda associada ao conceito de entropia que será tratado nas seções seguintes.

2.3 A entropia de Shannon: medida de informação e incerteza

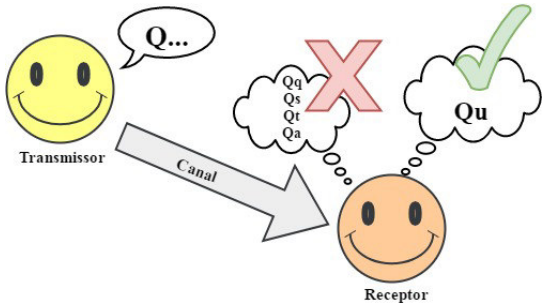
Como já mencionado, a entropia como forma de medir a informação contida em uma mensagem é o requisito fundamental para a aplicação dos demais teoremas de codificação de fonte e canal, que visam a comunicação entre máquinas. No entanto, devido à amplitude e plasticidade da informação, seu uso é adotado em outras áreas e com diferentes aplicações. Por isso, de acordo com Tribus e McIrvine (1971), a entropia de Shannon é considerada a medida fundamental para a ciência da informação, assim como o teorema de Pitágoras é para a geometria.

O "ponto de partida" de Shannon foi encontrar uma forma matemática de medir o quanto de informação existe na transmissão de uma mensagem de um ponto a outro, denominando-a entropia. Tal proposta baseava-se em conceitos estatísticos, expressando a ideia de que o aumento da probabilidade do próximo símbolo diminuiria o tamanho da informação (CHERRY, 1951). Assim sendo, há uma relação muito forte entre o conceito de entropia de Shannon e o conceito de incerteza.

Por isso, a entropia pode ser definida como a quantidade de incerteza que há em uma mensagem, a qual diminui à medida que os símbolos são transmitidos, ou seja, à medida que a mensagem vai sendo conhecida, tendo-se então a informação, que pode ser vista como redução da incerteza (SHANNON, 1948). Tal definição tem uma relação direta com o conceito de *negentropia* de Norbert Wiener, que será abordada com maior ênfase nos tópicos seguintes desta dissertação.

Como exemplo didático, ao utilizar como idioma a nossa língua portuguesa e ao transmitir como símbolo a letra “q”, a probabilidade do próximo símbolo ser a letra “u” é maior que a de ser qualquer outro símbolo, enquanto que a probabilidade de ser novamente a letra “q” é praticamente nula. A Figura 2.6 ilustra este exemplo didático.

Figura 2.6 - Exemplo didático da relação entre entropia e incerteza baseada na estatística



Fonte: Produção do próprio autor.

Shannon define que a entropia pode ser calculada através da soma das probabilidades de ocorrência de cada símbolo pela expressão $\sum p_i = 1 = 100\%$, onde p_i representa a probabilidade do i-ésimo símbolo que compõe a mensagem. Como a proposta de Shannon entende que estes símbolos devem ser representados através de sequências binárias, a utilização das propostas de Nyquist (1924) e Hartley (1928) torna-se necessária, prática e evidente. Tal proposta consistia em representar símbolos de um alfabeto através de um logaritmo de acordo com suas respectivas unidades de informação.

De acordo com Moser e Chen (2012), a entropia proposta por Shannon é obtida pela média das medidas de Hartley (1928). Sendo assim, Shannon (1948) apresenta a equação 2.1 como forma de computar a medida estatística desta incerteza com base nas probabilidades de representação da informação.

$$H = \sum p_i \log_2 \frac{1}{p_i} = -\sum p_i \log_2 p_i \tag{2.1}$$

Onde:

H = Entropia, medida da taxa de informação [bit/símbolo][9].

p_i = Probabilidade do i-ésimo resultado.

[9]Shannon também define H', uma variante simples de H, com a unidade [bit/segundo]

Assim, Shannon demonstrou que é possível fabricar uma mensagem tecnicamente através de escolhas, como um compositor escolhe letras, um músico escolhe notas e um pintor escolhe cores (WOODWARD; DAVIES, 1952). Portanto, é através da entropia que o trabalho de um engenheiro de telecomunicações reconfigura-se como o de conciliar a quantidade de seleções necessárias para a fabricação de unidades de informação (bits) na unidade de tempo disponível, segundo os limites físicos do canal de transmissão. Tal assunto é explorado através do Teorema da Capacidade de Canal, a ser explicado ainda neste capítulo.

2.3.1 "Inventores" da entropia na Teoria da Informação

A existência de confusões tal qual a abordada na seção 2.2.2 que aponta para Shannon-Weaver como "inventores" da fórmula da entropia quando o correto seria somente Shannon, é muito comum e estende-se a outros cientistas como Wiener, Hartley, Nyquist e Turing, por exemplo. Existem várias situações onde as importantes contribuições dadas a este campo científico por Wiener, Hartley e, de maneira menos recorrente por Nyquist, induzem a uma interpretação errada. O próprio Shannon reconhece estas valiosas contribuições, citando-as em seu artigo seminal.

De acordo com Gappmair (1999), Wiener definiu "informação = entropia" antes de Shannon, porém não tinha um conceito formal. A mesma premissa é abordada por Campbell (1982), onde há uma nota de rodapé com um relato da filha de Wiener afirmando que seu pai seria o idealizador da relação entre informação e entropia.

A proximidade entre Shannon e Wiener em termos de pesquisa, e pessoal também, representa uma confusão. É nitidamente próxima quando observada pelo contexto histórico em que viveram. De acordo com (GUIZZO, 2003), durante a segunda guerra mundial os Laboratórios Bell receberam a tarefa militar de melhorar o sistema de defesa anti-aérea. Tal projeto era chefiado por Vannevar Bush, orientador de Shannon tanto no mestrado quanto no doutorado. Nesta tarefa, Wiener elaborou um projeto anti-aéreo capaz de predizer, com base na estatística, as coordenadas de um avião. Este trabalho não foi publicado devido à guerra, porém ficou muito conhecido internamente e era chamado de "Perigo Amarelo" devido à cor amarela de sua capa. Outro motivo que justifica a aproximação de pesquisa entre os dois reside no fato de Shannon, enquanto estudante no MIT, ter feito um curso com Wiener sobre a teoria de Fourier.

No ano de 1948, após a publicação da Teoria Matemática da Comunicação de Shannon, Wiener consolidou-se como o pai da Cibernética com a publicação de um de seus principais

trabalhos: um livro intitulado "*Cybernetics: Or Control and Communication in the Animal and the Machine*", considerado como o alicerce desse campo de pesquisa. Neste trabalho, há um capítulo em que Wiener dedica-se a discussões matemáticas mais específicas, sobre "Séries temporais, informação e comunicação" em que são abordados conceitos que tratam da transmissão ou gravação de um sinal analógico variável através da sequência de amostras numéricas. Neste mesmo capítulo Wiener também faz uma análise da relação entre largura de banda, ruído e capacidade de informação, onde cita o trabalho de Shannon (WIENER, 1948; WIENER, 1961).

A aproximação entre Shannon e Wiener deu-se por correspondência, em outubro de 1948. Shannon escreveu a Wiener dizendo que havia lido *Cybernetics* e achado interessante a estreita relação do livro de Wiener com seu trabalho. No final da carta, Shannon pediu para que Wiener fizesse qualquer comentário sobre sua Teoria Matemática da Comunicação. Wiener respondeu a carta, agradecendo a Shannon pelo interesse e dizendo que também valorizou o trabalho de Shannon (GUIZZO, 2003).

A relação entre informação e incerteza também é abordada por Wiener em suas obras. Neste caso, a função da informação como a de reduzir a incerteza, é definida pelo termo negentropia, o inverso da entropia. No entanto, Wiener aborda este conceito de informação *versus* incerteza aplicado a problemas mais gerais de comunicação e controle, abordando desde a comunicação animal até mais especificamente a humana, inclusive aspectos biológicos, característicos do campo da Cibernética (WIENER, 1948; WIENER, 1961; WIENER, 1973).

No entanto, quando a informação é associada à incerteza e considerada um problema estatístico, a contribuição de Wiener na concepção deste conceito é reconhecida e referenciada pelo próprio Shannon (1948) em uma nota de rodapé na página 34 de sua teoria, bem como por outros pesquisadores como por exemplo: Reza (1961), Campbell (1982), Mandrekar e Masani (1997) e Seising (2009). De acordo com (GUIZZO, 2003), no MIT, pouco tempo depois da guerra, Wiener entrou no escritório do Prof. Dr. Robert Fano e declarou: "A informação é entropia." e saiu sem dar mais informações.

Através desta leitura, é possível observar que Wiener contribuiu com o conceito de entropia aplicado à Teoria da Informação com uma grande força teórica, porém muito abrangente, abarcando além da comunicação, campos muito densos como biologia e sociologia. Por outro lado, Shannon foi mais pragmático, focou em efetividade da Teoria da Informação aplicada à comunicação entre máquinas.

Apesar das ideias de Shannon e Wiener serem semelhantes, trabalhavam com problemáticas diferentes. Wiener preocupava-se em filtrar o ruído de um sinal recebido e Shannon em conviver com o ruído em uma transmissão de sinal (GUIZZO, 2003). De acordo com Pierce (2012), expressões semelhantes aparecem em ambos os trabalhos, porém destaca a singular interpretação de Shannon. Também aponta que o nome de Wiener passou a ser associado ao campo de extração de sinais de um determinado conjunto de ruídos enquanto que o nome de Shannon passou a ser associado a questões de codificação que possibilitem a transmissão com precisão e rapidez na presença de ruído.

Uma relação semelhante é entre Shannon e Alan Turing (1912-1954). Turing é conhecido por decifrar o "Enigma", uma máquina de criptografia e descriptografia de códigos de guerra usada pelos alemães durante a segunda guerra mundial, e também por conceber suas "máquinas de Turing". Neste caso, além de estar no mesmo contexto histórico de Shannon, também há uma proximidade de ideias, principalmente relacionadas à proposta de encontrar uma forma de mediar a informação e de usar a codificação para a proteção da informação. No entanto, com problemáticas diferentes, onde na teoria de Shannon a proteção é contra ruídos e Turing aborda a proteção contra a espionagem (GUIZZO, 2003).

Outra situação muito comum refere-se à citação Shannon-Hartley. Ao buscar nas principais bases científicas de indexação de periódicos, há muitas menções da fórmula da entropia como sendo a fórmula de "Shannon-Hartley". Uma busca simples pela palavra-chave "Shannon-Hartley" na base de periódicos da CAPES resultou em mais de 400 periódicos que usam este termo (CAPES, 2017). Alguns destes trabalhos são de outros campos que utilizam conceitos probabilísticos da entropia como metodologia para suas pesquisas científicas. Outros são até mesmo das áreas de comunicação, dentre eles pode-se destacar: Engenharia de comunicação; Computação; Biologia. Desta forma, surge mais um questionamento interessante: qual é a contribuição de Ralph Vinton Lyon Hartley (1988-1970) para a fórmula de entropia de Shannon?

Dois importantes estudos da década de 20 fundamentaram a proposta de entropia de Shannon: Nyquist (1924) e Hartley (1928). Tais estudos jamais foram ocultados, pelo contrário foram reconhecidos pelo próprio Shannon e considerados por ele um dos fatores de motivação a pensar a informação como um processo aleatório (ELLERSICK, 1984). Historicamente, antes de relatar as contribuições de Hartley, é interessante abordar as de Nyquist.

Harry Nyquist (1889-1976) fazia estudos sobre fatores que afetavam a velocidade de telégrafos e propôs ideias importantes como a de que o tempo pode ser discreto e de que é possível amostrar sinais analógicos de forma digital através do Teorema da Amostragem. Tal teorema é amplamente utilizado nos dias de hoje para processamento digital de sinais.

Nyquist utilizou a análise de Fourier para estudar a transmissão de sinais em telégrafo. Através desta técnica matemática foi possível a decomposição de um sinal complicado em uma soma de componentes mais simples. Simplificação esta que o fez perceber como esses componentes afetam a velocidade da transmissão. Ao aprofundar seus estudos sobre a aproximação de um sinal contínuo com uma série de componentes discretos, identificou a possibilidade de reconstruir o sinal, outra técnica amplamente explorada pela área de processamento digital de sinais (GUIZZO, 2003; NALON, 2009). Outra contribuição importante de Nyquist, de acordo com Lundheim (2002) e Guizzo (2003), foi a ideia de que a velocidade da transmissão sempre era limitada pela largura da banda do canal, expressa de forma matemática como:

$$W = K \cdot \log M \tag{2.2}$$

onde:

W = velocidade da transmissão

K = constante

M = número de símbolos do alfabeto

Ainda outra contribuição muito importante à Teoria da Informação foi dada quatro anos após os trabalhos de Nyquist. Trata-se do conceito abordado por Hartley (1928) que propunha uma forma de medir a quantidade de informação através de um logaritmo composto pelo número de mensagens possíveis (MASSEY, 1984; PIERCE, 1973). A contribuição de Hartley, de acordo com Lundheim (2002) e Massey (1984), foi importante para que Shannon reconhecesse que essa quantidade poderia ser medida matematicamente através da probabilidade da ocorrência dos símbolos em uma transmissão. De acordo com Guizzo (2003), Shannon conheceu e interessou-se pelo trabalho de Hartley quando ainda era aluno em Michigan, antes de ingressar no Bell Labs.

Hartley observou que quanto mais palavras tivesse uma linguagem, mais escolhas seriam possíveis de serem feitas e mais informações poderiam ser transmitidas. Desta forma, Hartley buscava medir a informação com base nesta liberdade de escolha (GUIZZO, 2003).

Para ilustrar essa relação, supõe-se a construção da palavra “DADO” usando um alfabeto de 26 símbolos, neste caso letras. A palavra “DADO” é formada por 4 símbolos deste alfabeto, portanto essa relação pode ser observada pelas etapas:

1. “D ? ? ?”: Para a escolha do primeiro símbolo há 26 possibilidades (26^1).
2. “D A ? ?”: Considerando o tamanho de 2 símbolos, haveriam 276 possibilidades (26^2), sendo 26 para o primeiro e outras 26 para o segundo símbolo.
3. “D A D ?”: Com o tamanho de 3 símbolos, seriam 17.576 possibilidades (26^3).
4. “D A D O”: Por fim, considerando o tamanho de 4 símbolos, seriam 456.976 possibilidades (26^4).

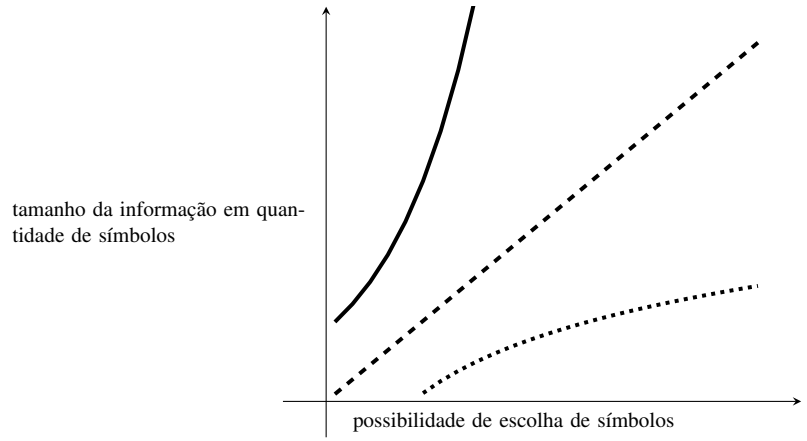
Desta forma, Hartley percebeu que o tamanho da informação transmitida cresce muito lentamente em relação ao número de possibilidades, que cresce exponencialmente. Isso é um grande problema, pois quanto maior a informação, muito maior seriam as possibilidades, por isso a busca deveria se iniciar por uma solução que fizesse as possibilidades crescerem linearmente ao invés de exponencialmente. Sendo assim, a solução matemática para este problema viria através da função matemática que é o “inverso” da exponenciação, a função logaritmo, pois quando um logaritmo é aplicado a uma curva exponencial, a curva torna-se uma linha reta (GUIZZO, 2003). A Figura 2.7 representa essa técnica matemática. Como resultado desta derivação matemática, (HARTLEY, 1928) apresentou a expressão:

$$H = \log S^n \tag{2.3}$$

onde:
H = Medida da informação
S = Número de símbolos possíveis de um alfabeto
n = Número de símbolos de uma transmissão

Desta forma, é possível fazer uma conexão entre a proposta de Hartley (1928) com a visão de Reza (1961) de os sistemas de comunicação serem considerados naturalmente estatísticos. Por isso, ainda de acordo com Reza (1961), a performance destes sistemas nunca pode ser descrita em um sentido determinista, apenas em termos estatísticos. Uma fonte ou receptor é um dispositivo que seleciona e transmite sequências e símbolos de um determinado alfabeto.

Figura 2.7 - Representação gráfica da técnica matemática usada por Hartley



Fonte: Produção do próprio autor. A técnica usada por Hartley consistiu na aplicação de uma função logarítmica representada pela linha pontilhada a uma função exponencial representada pela linha contínua, resultando em uma linha reta representada pela linha tracejada. Desta forma, demonstra-se a proposta de linearizar o crescimento das possibilidades em relação ao tamanho das informações

Cada seleção é feita aleatoriamente, embora esta seleção possa ser baseada em alguma regra estatística. O canal apenas transmite os símbolos de entrada para o receptor.

No entanto, a proposta de Hartley atendia somente processos discretos. Porém, para formas de comunicação contínuas, como telefone, rádio ou televisão, poderia ser usada a técnica de "quantização" ou Teorema da Amostragem de Nyquist (1924). Estas técnicas foram combinadas e implementadas em uma das primeiras transmissões digitais: um projeto conduzido pela Bell Labs com fins militares denominado "Sistema X". Este projeto visava proteger conversas telefônicas com criptografia entre Washington e Londres durante a segunda guerra mundial (GUIZZO, 2003).

De acordo com Guizzo (2003), durante a segunda guerra mundial, a técnica de criptografia já era amplamente utilizada em transmissões discretas feitas por telégrafos. Essa técnica consistia na implementação de uma chave de codificação que tornasse os dados interceptados ilegíveis, sendo que somente o receptor autorizado detinha a chave para realizar a decodificação. O grande desafio da equipe que trabalhava no Sistema X era a dificuldade de implementar códigos de criptografia para canais contínuos de comunicação como o telefone, pois tais canais usavam unidades de medida representadas através de sinais infinitesimais. Como já mencionado, derivados da combinação das técnicas de Nyquist (1924) e Hartley (1928), os valores numéricos eram convertidos na base 2. Desta forma, uma determinada altura de onda do sinal

contínuo seria transformada em um código binário. Assim, representando os 0's e 1's pela ausência ou presença de eletricidade, os canais discretos passaram a ser usados para a transmissão de dados representados de forma discreta. Após o término da guerra, este método foi amplamente estudado e desenvolvido no Bell Labs, dando origem ao PCM (Pulse Code Modulation) e amplamente divulgado através da publicação do artigo de Oliver, Pierce e Shannon (1948).

Outra importante contribuição dada por Nyquist (1924) à Teoria de Informação foi a ideia de que o alfabeto de uma determinada linguagem pudesse ser reduzido à apenas dois símbolos, que pudessem ser combinados com probabilidades em um logaritmo para representar as mensagens. De acordo com Cherry (1951) esse conceito de comunicação em que a informação é codificada a partir de composições formadas por apenas dois símbolos é antigo e pode ser encontrado em diversos exemplos usados por tribos primitivas como sinais de fumaça, tambores no Congo e inclusive o próprio Código Morse.

No entanto, esse conceito é explicado duas décadas após por Shannon (1948), quando observou que o logaritmo na base 2 desempenharia um papel eminente nesse contexto e batizou, por sugestão de J. W. Tukey, a unidade de informação como "dígito binário" (*Bit - Binary Digit* em inglês), fortalecendo então a tese de que a comunicação digital, baseada no *bit* seria mais eficiente que a analógica (GALLAGER, 2001; GAPPMAIR, 1999; MASSEY, 1984). De acordo com Guizzo (2003, p. 30, tradução nossa)

> A história diz que, um dia durante o almoço, alguns pesquisadores do Bell Labs estavam pensando em um termo melhor para dígito binário. Binit? Ou talvez bigit? Quando John Tukey, um dos homens da mesa colocou fim à discussão: a escolha melhor e óbvia, disse ele, é *bit*"[10].

Devido à condução das técnicas já existentes para o sistema de transmissão digital idealizado por Shannon, o alfabeto seria sintetizado em apenas dois símbolos: 0 e 1. Deste modo, a base do logaritmo de Hartley (1928) passaria a ser 2. Neste contexto, dois exemplos desta prática e evidência são apontados por Fano (1950). O primeiro trata-se de uma informação com 2 dígitos binários e o segundo de uma informação com 3. A Tabela 2.1 representa o conjunto de possibilidades de representação de uma informação de dois dígitos e a Tabela 2.2 uma informação de 3 dígitos.

Com base nas Tabelas 2.1 e 2.2, Fano (1950) nos aponta que quando tem-se conhecimento do "tamanho" de uma informação, nestes exemplos expressos respectivamente pela

[10] *"The story goes that, one day during lunch, some Bell Labs researchers were thinking of a better term for binary digit. What about binit? Or maybe bigit? John Tukey, one of the men at the table put an end to the discussion: the best and obvious choice, he said, was bit."*

Tabela 2.1 - Possibilidades de representação de uma informação de 2 *bits* através de uma tabela verdade

1 Dígito	2 Dígito
0	0
0	1
1	0
1	1

Fonte: Fano (1950)

Tabela 2.2 - Possibilidades de representação de uma informação de 3 *bits* através de uma tabela verdade

1 Dígito	2 Dígito	3 Dígito
0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1

Fonte: Fano (1950)

quantidade de dígitos x, é possível através da expressão 2^x determinar a quantidade de linhas da tabela verdade, na qual cada linha representa uma possibilidade de representação através dos dígitos binários. Desta forma, no primeiro exemplo tem-se $2^2 = 4$, portanto a Tabela 2.1 tem 4 linhas e no segundo exemplo $2^3 = 8$ resulta-se na Tabela 2.2 com 8 linhas.

No entanto, Fano (1950) e Moser e Chen (2012) apontam a ideia contrária trazida por Shannon (1948): de determinar o tamanho da informação a partir das possibilidades de representação (quantidade de linhas da tabela verdade, nestes exemplos). Neste caso, baseado na ideia de Hartley (1928), um logaritmo de base 2 do número de possibilidades n seria a forma matemática de representar o tamanho desta informação $H_n = \log_2 n$. Nos exemplos anteriores, considerando cada linha da Tabela 2.1 como uma possibilidade de representação, tem-se $H_n = \log_2(4) = 2$ *bits* e para a Tabela 2.2 de 8 linhas tem-se $H_n = \log_2(8) = 3$ *bits*.

Desta forma, a função $\log_2(x)$ teria o papel de medir o tamanho de um símbolo. Seria muito fácil deduzir que o somatório dos logaritmos destes símbolos resultaria no tamanho da

informação. No entanto havia um novo problema: a incerteza, determinada pelas probabilidades de ocorrência dos símbolos em um determinado idioma. No exemplo utilizado anteriormente da palavra "DADO', a construção desta palavra com base neste novo problema seria:

1. "D ? ? ?": Para a escolha do primeiro símbolo há 26 possibilidades (26^1). Portanto a incerteza é extremamente alta.

2. "D A ? ?": Considerando o tamanho de 2 símbolos, haveria 276 possibilidades (26^2). No entanto, na língua portuguesa, a probabilidade de a segunda letra precedente a uma letra D ser uma vogal é muito maior do que a de ser uma consoante. Neste mesmo caso, a probabilidade de ser novamente a letra D é praticamente inexistente. Portanto, nota-se que a partir do segundo símbolo, a incerteza diminui.

Desta forma, Shannon observou a relação entre o tamanho da informação com o tamanho da incerteza. À medida que os símbolos vão surgindo, a incerteza diminui. A forma que encontrou para medir essa incerteza é associar as funções logarítmicas com probabilidades, apresentando a fórmula 2.1 que foi batizada por ele de entropia (GAPPMAIR, 1999).

De acordo com (NEBEKER, 1998b), Shannon foi influenciado por Hartley e por Wiener quanto à natureza estatística da comunicação. Portanto, é possível deduzir-se, como é clássico na ciência, que a fórmula da entropia de Shannon foi produzida por Shannon levando em conta conceitos trabalhados por outros cientistas em períodos anteriores a ele. Isso não tira de Shannon o mérito de sua descoberta, haja vista que sua abordagem para o modelo matemático para comunicação foi inédita.

O aprimoramento das ideias já existentes é um fluxo natural que contribui para elevar o nível da ciência. De acordo Pineda (2006), na ciência, o cientista que é capaz de criar, dar publicidade e demonstrar como sua criação afetará a vida das pessoas, é considerado "pai" desta criação. Por isso, considerando esse fluxo normal da elevação do nível da ciência e como a fórmula da entropia de Shannon tem afetado a vida das pessoas, seria interessante que o título de "pai" da Teoria da Informação estivesse atribuído somente a Shannon.

2.4 O conceito de entropia

A palavra entropia assim como a palavra probabilidade é usada em muitos sentidos diferentes, têm vários significados (HAMMING, 1991). Por isso, é interessante examiná-los,

pois devido à diversidade de leituras podem gerar equívocos. De acordo com Brissaud (2005) o conceito de entropia originou-se na física e tornou-se um conceito complexo que influencia diversos campos do conhecimento. A multiplicidade de leituras aumenta a partir da denominação de entropia, por Shannon, à medida de informação, escolha e incerteza. Se diz também que o conceito de entropia é relativo à existência humana, antrópico (VILLANI, 2008).

Considerando que a expansão da Teoria da Informação a outros campos de pesquisa se deve, fortemente, à plasticidade do conceito de entropia, então a compreensão de tal conceito implica na compreensão de suas ramificações. De onde vem o nome entropia? A entropia na física e na Teoria da Informação tem o mesmo significado? "H" é igual a "S"? Por que Shannon utilizou o nome entropia?

Por isso, entende-se que o resgate histórico do conceito de entropia torna possível o esclarecimento de leituras equivocadas e certamente contribui para a compreensão de problemas atuais. Desta forma, na seção seguinte inicia-se uma abordagem sobre a origem da palavra entropia. A seguir, expõe-se um resgate histórico sobre o uso da representação simbólica "S" e "H" para entropia. Finaliza-se apresentando alguns argumentos que podem clarear a origem da utilização da palavra entropia na Teoria da Informação. Após essas abordagens, faz-se uma exposição dos principais teoremas de Shannon envolvidos pelo conceito de entropia em Teoria da Informação.

2.4.1 A origem da palavra entropia

De acordo com Moser e Chen (2012, p. 1, tradução nossa), assim "como vários outros ramos da matemática, a Teoria da Informação tem uma origem física"[11]. O contexto histórico que motivou a origem da palavra entropia no século XIX iniciou-se no final do século XVIII, quando cientistas buscavam formas eficazes de utilizarem motores térmicos para substituírem o trabalho humano, berço da termodinâmica, em especial de sua segunda lei. Para isso, os estudos buscavam formas de determinar quanto calor seria necessário para realizar um determinado trabalho ou vice-versa. Esse contexto histórico foi dividido em duas fases, classificadas por visões diferentes: *Não atômica* e *atômica* (BEN-NAIM, 2010).

Na primeira fase, a não atômica, a visão corrente considerava os conceitos atômicos como hipotéticos, sustentados apenas por filósofos gregos de dois milênios atrás. Nesta fase, houve três contribuições muito importantes para a formulação da segunda lei da termodinâmica.

[11] *"Like several other branches of mathematics, information theory has a physical origin."*

A primeira iniciou-se com Nicolas Leonard Sadi Carnot (1796-1832), que trouxe a teoria sobre o conceito de fluido calórico, demonstrando que há um limite superior para a eficiência do motor térmico e que esse limite depende exclusivamente das duas temperaturas do motor, a da sua fonte quente e da sua fonte fria, consequentemente que há um limite superior para o trabalho que pode-se obter de uma determinada quantidade de calor que flui das fonte quente à fonte fria (BEN-NAIM, 2010).

A segunda contribuição foi acrescentada por William Thompson (1775-1833), mais tarde conhecido como Lord Kelvin, através de uma descoberta importante que foi a escala de temperatura absoluta. Em uma primeira formulação para a segunda lei, Kelvin disse que não se pode converter totalmente energia térmica em trabalho, embora o reverso seja possível, isto é, o trabalho pode ser convertido completamente em energia térmica (BEN-NAIM, 2010).

A terceira e última contribuição da fase não atômica que praticamente sacramentou a segunda lei da termodinâmica foi a reformulação dos conceitos de Carnot e Kelvin por Rudolf Clausius (1822-1888). Clausius apresentou a ideia de que não pode haver um processo cujo único resultado seja um fluxo de energia de um corpo mais frio para o mais quente e então introduziu um novo termo, a entropia (BEN-NAIM, 2010).

Na teoria de Clausius, em qualquer processo espontâneo, ocorrendo em um sistema isolado, a entropia nunca diminui. De uma maneira simples, a entropia é entendida como calor perdido, ou seja, a parte de energia térmica que não poderá ser convertida em trabalho (BEN-NAIM, 2010). Embora Clausius tenha procurado uma palavra com origem em uma língua antiga, a grega, que pudesse expressar quantidades e que tivesse o mesmo significado em todas as línguas, a palavra entropia trouxe muito mistério. A palavra entropia é remetida pela etimologia à ideia de transformação, onde a primeira parte da palavra *en* lembra energia e a segunda parte lembra *tropos* que significa retorno, mudança (WEHRL, 1978). Clausius afirma sua escolha em:

> Eu considero melhor pedir emprestado termos para magnitudes importantes das línguas antigas, para que possam ser adotadas inalteradas em todas as línguas modernas, proponho então chamar a magnitude S de entropia do corpo, que vem da palavra grega τροπη, transformação [12] (CLAUSIUS, 1867, p. 357, tradução nossa).

Ainda de acordo com a opção de Clausius por essa palavra, ela deveria conotar a ideia de energia, como afirmado: "Tenho intencionalmente formado a palavra entropia de modo a ser

[12] *"[...] I hold it to be better to borrow terms for important magnitudes from the ancient languages, so that they may be adopted unchanged in all modern languages, I propose to call the magnitude S the entropy of the body, from the Greek word* τροπη*, transformation.'*

o mais parecido possível com a palavra energia [...]"[13] (CLAUSIUS, 1867, p. 357, tradução nossa).

Referente a esta primeira fase classificada por Ben-Naim (2010) como não atômica, Hamming (1991) a chama de "Termodinâmica Clássica" e reconhece que é onde aparentemente surge o termo entropia. Também complementa mencionando que nesta fase a grande preocupação era a tratativa das variáveis macroscópicas de estado que caracterizam todo sistema físico: pressão, volume e temperatura. Destas variáveis diretamente mensuráveis, dependem várias outras variáveis internas do sistema como: entropia, entalpia e energia interna.

A expressão matemática da 2ª lei da termodinâmica é dada por:

$$dH = \frac{dQ}{T} \tag{2.4}$$

Onde:

dH = Alteração na Entropia

dQ = Quantidade de Calor Transferido

T = Temperatura Absoluta

No final do século XIX, inicia-se a segunda fase da termodinâmica: a Atômica. Neste contexto, a física já estava dividida em mecânica, eletromagnetismo e termodinâmica. Começava a aceitação do modelo atomístico da matéria, porém o novo desafio passou a ser o de lidar com um grande número de partículas. Para tal, foram necessários métodos estatísticos que consequentemente exigem ferramentas de uma teoria de probabilidades. Neste cenário, grandes cientistas se destacaram, entre eles James Clerk Maxwell. Ele introduziu conceitos estatísticos na termodinâmica e criou a "Teoria Cinética dos Gases". Nesta teoria, as probabilidades foram utilizadas apenas como ferramentas auxiliares para calcular quantidades médias tais como o volume, a pressão e temperatura. No entanto, quantidades médias e probabilidades não eram bem aceitas por físicos mais conservadores (BEN-NAIM, 2010).

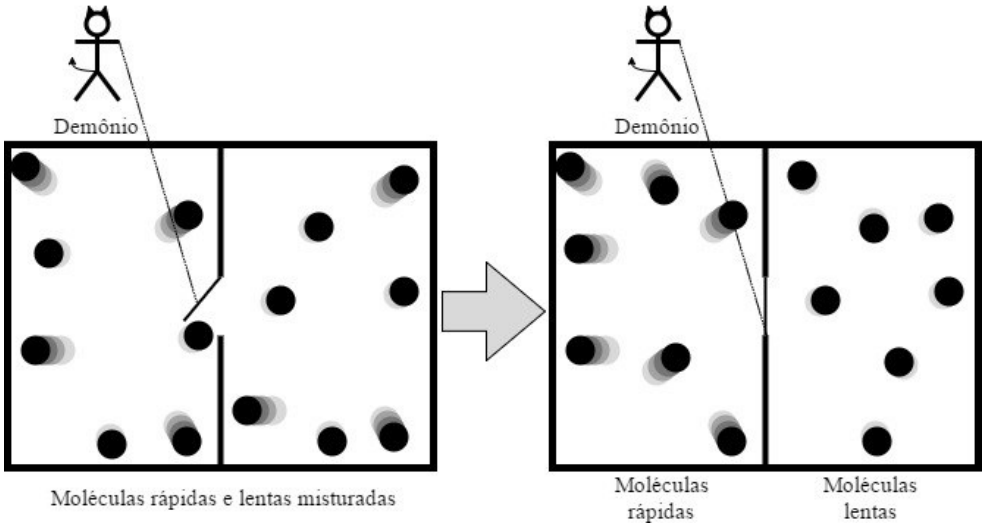
Outro trabalho muito importante de Maxwell, que merece destaque por estar diretamente ligado à entropia, foi uma proposta para contrariar a segunda lei da termodinâmica que afirma que num sistema isolado a entropia nunca diminui. Nesta proposta, ele criou um problema imaginário conhecido como o "demônio de Maxwell". A partir de um personagem

[13] *"I have intentionally formed the word entropy so as to be as similar as possible to the word energy[...]"*

imaginário, chamado de demônio, Maxwell (1871) demonstrou ser possível, através da reorganização das moléculas, com ajuda desse demônio, diminuir a entropia de um sistema isolado.

Neste problema, Maxwell (1871), propõe que se imagine uma caixa isolada com gás em seu interior. Esta caixa é dividida ao meio por um pequeno portão. O demônio era tido como o "porteiro" deste cenário, de modo que sua função fosse a de observar as moléculas do gás presentes em ambos os lados da caixa. Inicialmente, o gás possuía a mesma temperatura em ambos os lados. A temperatura do gás é uma medida da velocidade média de suas moléculas. A ideia de entropia era associada a ordem e desordem, ou seja, quanto mais agitadas as moléculas, maior desordem, maior temperatura e consequentemente maior entropia. As velocidades destas moléculas são diferentes. Neste contexto, a função do demônio era observar as velocidades das moléculas e separar, em cada lado da caixa, as mais rápidas das mais lentas. Desta forma, após a separação, o lado da caixa que contém as moléculas mais rápidas ficaria mais quente e o outro lado, com as mais lentas, mais frio (GUIZZO, 2003). A Figura 2.8 ilustra essa problemática trazida por Maxwell. Nesta nova situação, pode-se mostrar que a entropia total da caixa é menor que na situação inicial.

Figura 2.8 - Representação ilustrativa do "demônio de Maxwell"



Fonte: Adaptada de Eler (2017).

A partir do problema de Maxwell, Ludwig Boltzmann propôs uma reformulação da segunda lei, baseada em probabilidade. Na reformulação de Boltzmann, fenômenos que eram considerados impossíveis (espontaneamente irreversíveis) passam a serem considerados altamente improváveis. Esta reformulação probabilística não foi aceita rapidamente, pois para

a física, uma lei deve ser absoluta e exceções não são permitidas, por isso probabilidades não eram bem vistas (BEN-NAIM, 2010). A existência dos átomos ainda era muito questionada e criticada, consequentemente a teoria de Boltzmann também o foi (PINEDA, 2006).

Dentre as várias contribuições de Boltzmann utilizando a abordagem estatística da termodinâmica, pode-se citar a relação entre a temperatura e o movimento médio das moléculas, provando a lei de distribuição de velocidades de Maxwell. Boltzmann também propôs um mecanismo para provar que a desordem molecular é causada por essa distribuição de velocidades, ao qual chamou de "Teorema H" e definiu a entropia como uma medida estatística dessa desordem (PINEDA, 2006).

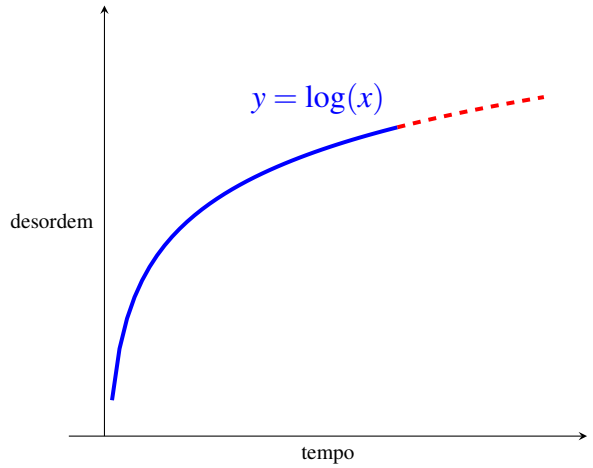
Boltzmann afirmou que a distribuição de velocidades das moléculas segue uma estatística hoje conhecida como estatística de Maxwell-Boltzmann. Partindo da premissa que todas as moléculas de um gás isolado tenham inicialmente a mesma velocidade, com o decorrer do tempo ocorrerão colisões entre as moléculas, que provocarão alterações em suas velocidades. Desta forma, algumas ficarão mais rápidas e outras mais lentas. Assim, o sistema parte de uma estrutura inicial "ordenada" para atingir o máximo de "desordem". Quando o máximo de desordem é atingido, o sistema entra em uma fase de estabilidade, em que as colisões entre as moléculas não interferem mais na velocidade média das mesmas (PINEDA, 2006).

Associando velocidade à temperatura, Boltzmann sugere que a temperatura do sistema se estabiliza quando atinge o máximo de desorganização. Sendo assim, se diz que o retorno ao estado inicial ordenado não é impossível, mas é extremamente improvável. Um exemplo apontado seria a mistura de dois gases em um recipiente de 100 ml, quando atingisse seu grau máximo de desordem, demoraria cerca de $10^{10000000000}$ anos para que a configuração das moléculas se tornasse idêntica à inicial. De forma prática, isso significaria nunca. Foi através do "Teorema-H" que Boltzmann provou estas ideias (PINEDA, 2006).

Desta forma, uma função logarítmica seria ideal para expressar o crescimento da desordem entre as moléculas no decorrer do tempo até uma estabilização do valor médio das grandezas físicas macroscópicas, momento em que a desordem atingiria um máximo estável. Tal representação é esboçada na Figura 2.9. No entanto, ao considerar a utilização de uma mistura de gases diferentes, haveria moléculas de tamanhos diferentes: m e $m1$. Por isso, para este caso haveria 3 possibilidades diferentes de colisão: as moléculas m entre si, as moléculas $m1$ entre si e as moléculas m com as $m1$. Boltzmann atribuiu a probabilidade de $\frac{1}{4}$ para as colisões entre as mesmas moléculas e $\frac{1}{2}$ para as colisões entre moléculas diferentes. Desta forma,

foi necessário associar probabilidades a funções logarítmicas, associação esta que também é o fundamento da entropia na Teoria de Informação (PINEDA, 2006).

Figura 2.9 - Representação análoga da desordem de moléculas de gases em relação ao tempo



Fonte: Produção do próprio autor. Ilustra-se através deste gráfico da função logarítmica $y = \log(x)$ uma analogia entre o teorema de Boltzmann para o crescimento da desordem das moléculas de gases em relação ao tempo. A linha contínua expressa o crescimento desta desordem em função do tempo e a linha tracejada aponta para a tendência de relativa estabilização do crescimento no decorrer do tempo.

Ainda no contexto da termodinâmica onde discute-se a estatística aplicada a moléculas de gases, surge a primeira conexão entre entropia e informação. Ela é feita muito antes de Shannon, através do físico Leo Szilard, em 1929. A proposta de Szilard (1964)[14] consistiu na "exorcização" do "demônio de Maxwell", partindo da problemática de como o demônio teria conhecimento da velocidade das moléculas para separá-las em cada lado da caixa. Desta forma, Szilard concluiu que a identificação desta velocidade viria através de uma informação, neste caso a velocidade da molécula, que poderia ser aferida somente através da interação demônio-molécula (GUIZZO, 2003; PINEDA, 2006). Desta forma, pode-se entender que, neste contexto histórico, estatística, matemática e física já estavam "maduros" suficientemente para que um novo conceito começasse a surgir. Szilard captou esse amadurecimento científico.

O novo problema apontado por Szilard era que esta interação, visando diminuir a entropia, causaria desordem ao sistema aumentando a entropia geral. Desta forma, a redução da temperatura estaria associada ao ganho de informação. Imaginariamente, o demônio teria que perguntar para cada molécula o valor de sua velocidade ou usar um instrumento para medi-la

[14] Em 1964 a obra original de 1929 foi traduzida e republicada. Nas referências, é apontada a versão republicada

uma a uma. Pode-se então imaginar a "bagunça" que isso causaria. Logo, maior desordem, maior será a entropia e maior a temperatura. Desta forma, a segunda lei da termodinâmica era comprovada novamente (GUIZZO, 2003; PINEDA, 2006). Ainda assim, a questão da relação desordem/entropia está longe de ser tão clara ou simples. Isto pode muito bem estar relacionado ao fato de que o próprio conceito de ordem e desordem é essencialmente antrópico, conforme apontado anteriormente (HOSSENFELDER, 2018).

2.4.2 A palavra entropia na Teoria da Informação

Uma "interrogação" persiste na Teoria da Informação e está ligada à utilização do nome entropia, de Clausius, por Shannon. Porque Shannon utilizou extamente o mesmo nome entropia para denominar sua medida de informação? Conforme já abordado na seção anterior, este nome já era amplamente utilizado na física e inclusive associado às medidas estocásticas com base em probabilidades. Na literatura científica há alguns trabalhos que citam que a palavra entropia fora sugerida a Shannon por John Von Neumann e outros apontam que Shannon teve iniciativa própria ao escolher a palavra entropia.

Em 1984, em uma entrevista concedida por Shannon ao Dr. Robert Price, quando questionado sobre a origem do nome entropia em Teoria da Informação, o entrevistador calçou sua pergunta afirmando que acreditava que a palavra entropia não fora sugerida por Von Neumman. Dentre os argumentos utilizados por Price, o primeiro foi que embora a palavra entropia estivesse associada à mecânica estatística, em 1945 no seu relatório sobre criptografia Shannon já usava esta palavra. Isto fez o entrevistador deduzir que já em 1945 Shannon tinha propósito de usar essa palavra e talvez tivesse consultado Von Neumann e demais outros colegas somente sobre a viabilidade deste uso. O entrevistador também apresentou um segundo argumento para não acreditar em um contato entre Shannon e Von Neumann em 1945. Em sua resposta, Shannon negou a participação de Von Neumann, dizendo: "Não, eu não acho que ele o fez. Tenho certeza de que isso não aconteceu entre Von Neumann e eu"[15] (ELLERSICK, 1984, p. 124, tradução nossa).

A primeira "interrogação" nesta entrevista está relacionada ao argumento do entrevistador de que Shannon teria usado a palavra entropia já em 1945, em seu relatório sobre criptografia. De acordo com (GUIZZO, 2003, p. 46, tradução nossa), há uma contradição, pois em 1945:

[15] *"No, I don't think he did. I'm quite sure that it did not happen between Von Neumann and me."*

> [...] Shannon considerava a fórmula logarítmica como uma medida de incerteza - não entropia. Mais tarde, a seção intitulada "Escolha, Informação e Incerteza" nesse relatório tornou-se "Escolha, Incerteza e Entropia" no artigo de 1948 [16].

De fato, este argumento procede ao analisar o relatório de 1945, no qual a palavra entropia aparece quando Shannon explica a característica de aleatoriedade de sua medida, até então chamada de incerteza, presente na mecânica estatística. Esta comparação fora realizada com o "Teorema-H" de Boltzmann, no qual Shannon conclui que "A maioria das fórmulas de entropia contém termos desse tipo[17]" (SHANNON, 1945, p. 10, tradução nossa). Logo que Shannon deduziu a sua fórmula para a medida da informação, ele observou que outras áreas já usavam a combinação da função logarítmica com probabilidades, dentre elas a mecânica estatística e a termodinâmica com o conceito de entropia (GUIZZO, 2003).

Portanto, é possível que Shannon, apesar de não ter chamado sua medida de informação de entropia em 1945, já a tenha como proposta de nome em 1948, atestando assim sua iniciativa própria na escolha deste nome. No entanto, um fato interessante a ser relevado que pode ser considerado uma segunda "interrogação" é que, de acordo com Gallager (2001), Shannon foi diagnosticado com Alzheimer na década de 80. Segundo Guizzo (2003), já na década de 90 Shannon tinha dificuldades de encontrar o caminho de volta para casa. Por isso, considerando que esta entrevista ocorreu em 1984, este argumento pode invalidar a proposição de Shannon de que Von Neumann não teve participação no projeto da unidade de medida da informação com o nome de entropia.

A premissa da literatura que contradiz a versão de independência de Shannon ao escolher o nome Entropia é trazida por (TRIBUS; MCIRVINE, 1971) e replicada em forma de citação em diversos artigos relacionados. Neste artigo, os autores afirmam que quando Shannon foi questionado por um deles (Tribus) em 1961 sobre a origem do nome entropia para sua medida de informação, Shannon respondeu que havia pensado tanto no nome informação quanto no nome incerteza, no entanto considerou que ambos eram usados com muita frequência, então ao conversar com John Von Neumann a respeito, ele sugeriu o nome entropia, justificando em primeiro lugar sua função de incerteza usada na mecânica estatística sob esse nome e, em segundo lugar, que ninguém saberia o que realmente era a entropia e, então em um debate, Shannon sempre teria vantagem.

[16] *"In his 1945 cryptography report, Shannon regarded the logarithmic formula as a measure of uncertainty-not entropy. But later, the section titled "Choice, Information and Uncertainty" in that report became "Choice, Uncertainty and Entropy" in the 1948 paper."*

[17] *"Most of the entropy formulas contain terms of this type."*

> "O que há num nome? No caso da medida de Shannon, o nome não foi acidental. Em 1961 um de nós (Tribus) perguntou a Shannon o que ele tinha pensado quando ele finalmente encontrou sua famosa medida. Shannon respondeu: 'Minha maior preocupação foi com o nome que daria a ela. Eu pensei em chamá-la de informação, mas a palavra foi excessivamente utilizada; assim, eu decidi chamá-la de incerteza. Quando eu discuti isso com John von Neumann, ele teve uma ideia melhor'. Von Neumann disse-me: 'Você deveria chamá-la de entropia, por duas razões. Em primeiro lugar, sua função de incerteza tem sido utilizada na mecânica estatística sob esse nome; assim, ela já está em um nome. Em segundo lugar, e mais importante, ninguém sabe o que realmente é entropia; assim, num debate você sempre tem vantagem." (TRIBUS; MCIRVINE, 1971, p. 180, tradução nossa)[18].

2.4.3 A entropia na Física e na Teoria da Informação

Apesar de a entropia de Shannon (Teoria da Informação) e Clausius (Física) serem representadas pelo mesmo nome, para Tribus e McIrvine (1971), em uma primeira vista, nada indica que tenham a mesma função, pois Clausius limitava-se à ideia de entropia relacionada à perda inevitável de calor. No entanto, ao considerar o posterior desenvolvimento deste conceito dentro da termodinâmica com Maxwell, Boltzmann e Szilard, a palavra entropia ganhou um significado com mais plasticidade, relacionado a eventos estocásticos, aleatoriedade, probabilidades e funções logarítmicas que podem ter influenciado as ideias de Shannon.

Uma analogia bem humorada sobre a relação da entropia de Shannon e a de Clausius, bem como a participação de John Von Neumann na criação do nome entropia é feita por Ben-Naim (2010, p. 14). Também, num recente livro, Ben-Naim trabalha a distinção entre os conceitos de entropia de Shannon, que ele chama de SMI (*Shannon's Measure of Information*), e de entropia da física (Clausius). Ele diz que a entropia da Física é um caso particular da entropia SMI de Shannon (BEN-NAIM, 2017).

Ao comparar as entropias da física e da Teoria da Informação, deve-se observar que

> O simples fato de que a mesma expressão matemática $-\sum p_i \log p_i$ ocorrer tanto na mecânica estatística quanto na teoria da informação, não estabelece, por si só, nenhuma conexão entre esses campos. Isso só pode ser feito ao encontrar novos pontos de vista para os quais a entropia da termodinâmica e

[18] *"What's in a name? In the case of Shannon's measure the naming was not accidental. In 1961 one of us (Tribus) asked Shannon what he had thought about when he had finally confirmed his famous measure. Shannon replied: "My greatest concern was what to call it. I thought of calling it 'information', but the word was overly used, so I decided to call it 'uncertainty'. When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, 'You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage."*

> na teoria da informação aparecem como o mesmo conceito[19] (JAYNES, 1957, p. 621, tradução nossa).

Na busca por estes novos pontos de vista que apontem semelhanças ou diferenças que possam clarear a confusão da ocorrência da entropia na relação entre estes diferentes campos, discute-se o conceito da irreversibilidade. De acordo com Pineda (2006), uma diferença associada a este fato na Teoria da Informação é que a entropia decresce, chegando a zero quanto há certeza total. Porém, na física, a entropia vista como a desordem de moléculas estabiliza, mas o decrescimento é extremamente improvável como apontava Boltzmann. Para Brillouin (1950), esta reversibilidade é apontada através do conceito de "negentropia". Desta forma, entende-se que uma informação que encontra-se em um ponto pode ser reproduzida em outro ponto somente através de um processo de codificação e, neste processo, a condição necessária é a reversibilidade. Neste caso, Brillouin (1950) cita o exemplo de traduzir um texto do inglês para japonês e ao traduzir de volta para o inglês, o texto deve ser idêntico ao original.

No entanto, mesmo para a reversibilidade, ainda há semelhança com a termodinâmica. Neste caso, a explicação vem através da degradação da informação. A informação pode ser classificada como absoluta ou distribuída. A informação absoluta é aquela produzida exclusivamente por humanos a partir da observação de fenômenos. No entanto, para que esta informação produzida chegue até a sociedade é necessária a redundância, ou seja, a reprodução desta informação absoluta produzida. Desta forma, a informação deixa de ser absoluta e passa a ser distribuída. Neste momento de distribuição, ocorrem as degradações da informação. Uma informação totalmente degradada não serviria para nada mais. Ao considerar novamente o exemplo de traduzir o texto do inglês para o japonês e depois de volta ao inglês, certamente haveria uma degradação neste processo e essa parte perdida de informação pode ser comparada à entropia de Clausius, como a parte de energia térmica que não poderia ser convertida em trabalho (BRILLOUIN, 1950).

Por isso, de acordo com Brillouin (1950), a 2ª Lei da Termodinâmica da Física e a Teoria da Informação estão conectadas através da ideia de entropia e de negentropia. Com base neste conceito de Brillouin (1950), na Tabela 2.3 mostra-se o exemplo de degradação de informação. Nesta tabela, a primeira linha contém uma parte do texto do parágrafo anterior em sua versão original, no idioma português. Na segunda linha, há uma versão do mesmo texto tradu-

[19] *"The mere fact that the same mathematical expression* $-\sum p_i \log p_i$ *occurs both in statistical mechanics and information theory does not in itself establish any connection between these fields. This can be done only by finding new viewpoints from which themodinamic entropy and information-theory entropy appear as the same concept"*

zido para o idioma inglês. Na terceira linha, há a tradução do texto para o português novamente. Observam-se algumas degradações na comparação da primeira com a terceira linhas. Os grifos representam tais degradações. Ambas as traduções foram realizadas através da ferramenta de tradução do Google (2017) como forma de representar a participação das máquinas no processo de distribuição da informação.

Tabela 2.3 - Exemplo de degradação da informação

Original	No entanto, mesmo para a reversibilidade ainda há semelhança com a termodinâmica. Neste caso, a explicação vem através da degradação da informação. A informação é classificada como absoluta ou distribuída. A informação absoluta é aquela produzida exclusivamente por humanos a partir da observação de fenômenos. No entanto, para que esta informação produzida chegue até a sociedade é necessária a redundância, ou seja, a reprodução desta informação absoluta produzida. Desta forma, a informação deixa de ser absoluta e passa a ser distribuída.
Português-inglês	However, even for reversibility there is still resemblance to thermodynamics. In this case, the explanation comes through the degradation of information. The information is classified as absolute or distributed. Absolute information is that produced exclusively by humans from the observation of phenomena. However, for this information produced to reach society, redundancy is needed, that is, the reproduction of this absolute information produced. In this way, the information stops being absolute and begins to be distributed.
Inglês-português	No entanto, mesmo para a reversibilidade, ainda existe semelhança com a termodinâmica. Neste caso, a explicação vem pela degradação da informação. A informação é classificada como absoluta ou distribuída. A informação absoluta é a produzida exclusivamente por seres humanos a partir da observação de fenômenos. No entanto, por essa informação produzida para alcançar a sociedade, é necessária a redundância, ou seja, a reprodução dessa informação absoluta produzida. Desta forma, a informação deixa de ser absoluta e começa a ser distribuída.

Fonte: Texto original produzido pelo próprio autor e traduções realizadas por Google (2017)

Outra diferença apontada entre as entropias é conceitual. Neste caso, a entropia da teoria da informação e a entropia da termodinâmica são conceitos inteiramente diferentes (JAYNES, 1963).

Uma outra comparação entre a entropia da física (termodinâmica) e da teoria da informação é feita de forma mais sintética por Pineda (2006). A Tabela 2.4 representa esta síntese comparativa adaptada.

Tabela 2.4 - Comparação entre o conceito de entropia na Física e na Teoria da Informação

	Física	**Teoria da Informação**
Autor	Boltzmann	Shannon
Objeto	Gases	Informação
Conceito	Desordem molecular	Quantidade de Informação
Medida	Distribuição de velocidades das moléculas	Quantidade de Informação
Variação	Aumenta à medida que o movimento das moléculas se torna aleatório	Aumenta a medida que a ocorrência dos símbolos se torna aleatória
Fórmula	$S = k \log W$	$-\sum P_i \log_2 P_i$

Fonte: Extraída e adaptada de: Pineda (2006, p. 88)

Além de Szilard, citado na seção 2.4.1, ter relacionado ainda na Física a entropia à informação, R. A. Fisher, em 1925 também utilizou, em termos técnicos, a palavra informação

em seu trabalho sobre teoria da estimação estatística, que segundo Kullback (1959), a medida da quantidade de informação fornecida por dados sobre parâmetros desconhecidos foi o primeiro uso de "informação" em matemática estatística. Desta forma, é possível a dedução de que a Teoria da Informação, em termos de conceitos e intuição, tem suas raízes matemáticas nos conceitos de desordem (ou entropia) da termodinâmica e na mecânica estatística.

Desta forma, a palavra entropia acaba se tornando um grande dilema, pois há argumentos que indicam grandes diferenças entre a entropia da Teoria da Informação e a entropia da física, mas também existem argumentos contrários que tornam evidentes algumas semelhanças. Diante do exposto torna-se difícil estabelecer qualquer prognóstico acerca da origem exata da escolha do nome entropia por Shannon. Torna-se difícil também indicar uma possível equivalência conceitual entre os conceitos de entropia na Teoria da Informação e na Termodinâmica.

2.4.4 A notação da entropia: "S" ou "H"?

Como pode ser observado de uma maneira mais evidente na Tabela 2.4, outra curiosidade interessante envolvendo o conceito de entropia, talvez reflexo de sua amplitude, está em sua notação. Existem trabalhos em que a fórmula da entropia é expressa pela letra H e casos em que é expressa pela letra S. Nos parágrafos seguintes são apresentados alguns argumentos e hipóteses, fazendo um resgate histórico e levando a uma reflexão sobre essa questão.

Historicamente, a palavra entropia de origem grega, escolhida por Clausius, é acompanhada da notação S. Em uma de suas obras, Clausius (1867, p. 357), usa essa notação: "[...] proponho chamar a magnitude S da entropia do corpo [...]"[20].

Um pouco mais tarde, o desenvolvimento do conceito de entropia por Maxwell e Boltzmann agrega-se à ideia original de Clausius, modificando a segunda lei da termodinâmica. No entanto, a entropia mantem-se preservada pela notação S, como podemos observar na fórmula 2.5 atribuída à Entropia de Boltzmann, (BEN-NAIM, 2010):

$$S = k \log W \tag{2.5}$$

Onde:

S = Entropia

[20] "*[...] I propose to call the magnitude S the entropy of the body [...].*"

k = Constante de Boltzmann que relaciona temperatura e energia de moléculas, pelo SI $k = 1,3806503 \cdot 10^{-23}\ \frac{J}{K}$ com K representando Kelvin (unidade de medida no SI para temperatura termodinâmica, expresso por $\frac{1}{273,16}$ da temperatura termodinâmica do ponto triplo da água) e J representando Joules no SI.

W = Probabilidade Termodinâmica

Com exceção de Clausius, na literatura científica revisada não foi encontrada menção à escolha da notação S para entropia[21]. No entanto, ao utilizar o argumento da origem grega da palavra entropia escolhida por Clausius e investigar o símbolo S no alfabeto grego, encontra-se o sigma, que de forma minúscula é simbolizado por σ e maiúscula por Σ. Ambos são muito usados em álgebra, inclusive, de acordo com Fiorentini, Miorin e Miguel (2016), o sigma foi um dos primeiros símbolos mais utilizados nas fases inicias da álgebra para representar incógnitas, fase denominada de sincopada. Esta fase estendeu-se até o início do século XVI. Quando expressa de forma minúscula representa ideias de transformações, incertezas, variações. Escrita de forma maiúscula é utilizada para representar séries de somas. Desta forma, a primeira hipótese que surge é se Clausius teria se baseado nestas tradições algébricas para atribuir a notação de S à entropia e então manter o estilo algébrico dos séculos anteriores.

Uma outra menção curiosa aparece relacionada à entropia quando Boltzmann (2003) apresenta o "Teorema-H", definindo como H o negativo da entropia (DIAS, 1994). Ou seja, neste contexto, a notação H é atribuída para o inverso da entropia, portanto $H = -S$ onde S ainda representa a entropia.

Um pouco mais adiante no decorrer histórico, a simbologia H é atribuída à entropia no trabalho de Burbury (1890) que faz um levantamento sobre alguns problemas existentes na *"Teoria Cinética dos Gases"* de Maxwell. Ao estudar a *"Lei de Distribuição de Velocidades"* de Maxwell, em particular quanto ao número de colisões de moléculas de gases em um cilindro, ele propõe uma fórmula de cálculo adaptada da fórmula de Boltzmann. De acordo com a nota de rodapé da página 6 da edição revisada de BRUSH (1966), a utilização da letra H para entropia se dá com Burbury neste artigo. Nesta nota de rodapé, BRUSH (1966) fundamenta sua afirmação no resumo da obra de Chapman (1937) que apontou uma confusão entre a letra H como oriunda da letra maiúscula grega *eta*.

[21] S. G. Brush escreve em uma nota de rodapé de seu livro, *Kinetic Theory*, que: "*Boltzmann himself originally used the letter E, and did not change to H until 1895; the first use of H for this quantity was apparently by S. H. Burbury, Phil. Mag. 30, 301 (1890)*" (BRUSH, 1966, p. 6).

Às vésperas da publicação do artigo de Shannon em 1948, Neumann (1944) em um trabalho publicado com o título de "*Proposal and analysis of a new numerical method for the treatment of hydrodynamical shock problems*" menciona a entropia, ainda que no significado termodinâmico e da mecânica estatística respeitando a notação de S, tradicionalmente usada na área.

A reviravolta das notações ocorre quando Shannon (1948) relaciona a entropia com a Teoria da Informação, apontando a letra "H" como notação. A relação entre o Teorema-H de Boltzmann e a entropia de Shannon notada com H é ressaltada pelo próprio Shannon em seu artigo. Reforçando esta mudança de notação, (HAMMING, 1991), aponta que ao lidar com "informação", a entropia é vista como padrão para a quantidade de informação e é geralmente rotulada pela notação $H(X)$ ou $H(P)$, onde X ou P representam uma variável aleatória discreta (HAMMING, 1991).

De acordo com Brookes (1956), assim como existem outras interessantes analogias entre sistemas termodinâmicos e de informação, foi por analogia a Boltzmann que Shannon chamou sua função de H.

Na própria literatura clássica há essa confusão de notações, ainda como exemplo podemos citar Hamming (1991) que ao se referir à "Lei da Conservação de Energia", cita a equação 2.4 utilizando dH em sua notação. Entretanto, ao se referir à mesma equação, Ben-Naim (2008, p. 7) utiliza a notação dS para entropia.

Como observa-se nos argumentos citados nesta seção, não há uma convenção clara sobre qual notação utilizar. No entanto, pode-se deduzir que há um senso comum na utilização da notação S para a área de física, mecânica estatística, mecânica quântica e correlatas, enquanto que a Teoria da Informação utiliza-se da notação H.

No entanto, uma relação entre a entropia notada com S e a notada com H é feita por Villani (2008), quando aponta a influência do Teorema-H de Boltzmann na Teoria da Informação. Ele faz a comparação:

$$-\sum f_j \log f_j$$

com

$$-\int f log f,$$

e explica que a entropia de Shannon-Boltzmann

$$S = -H = -\int flogf$$

quantifica (em escala logarítmica) quanto de informação existe em um sinal aleatório, ou em um idioma, considerando f como a densidade da distribuição do sinal. Em uma linguagem determinista, isto significa previsibilidade completa.

Portanto, não há na literatura um esclarecimento de qual notação é correta e existe o senso comum de H ser usado na Teoria da Informação e S na física.

2.5 A entropia e os principais teoremas de Shannon

Nas subseções seguintes expõem-se, de modo breve, os teoremas clássicos de Shannon: Teorema da Capacidade de Canal, Teorema de Codificação de Fonte e Teorema de Codificação de Canal (SHANNON, 1948).

2.5.1 Teorema de Capacidade de Canal

Além de Shannon, outros cientistas da época (e antes dele) trabalhavam na busca por melhorias no processo de comunicação. Alguns, antes de Shannon, já haviam proposto formas de medir a transmissão de um sinal. No entanto, o trabalho de Shannon se destaca por três grandes diferenciais: 1- Buscou uma solução considerando o ruído, pois outras descobertas semelhantes da época não incluíam o ruído (LUNDHEIM, 2002); 2- Provou, de forma concisa e prática que é possível transmitir dados de maneira confiável por um canal ruidoso (GALLAGER, 2001); 3- Apresentou ideias claras e concisas que nortearam a otimização do processo de comunicação através da codificação, seja para compressão ou para detecção e correção de erros (REZA, 1961; HUFFMAN, 1952; HAMMING, 1950).

Para tornar essa transmissão confiável, livre de erros, Shannon propõe um "teorema fundamental" para um canal discreto com ruído (GUIZZO, 2003), em que a quantidade de informação a ser transmitida seja sempre menor que a capacidade do canal, ou seja, a Entropia (H) jamais deve exceder a capacidade do canal, como representado na expressão 2.6:

$$C > H \tag{2.6}$$

Onde:

C = Capacidade do Canal

H = Entropia

No entanto, para permitir a comparação da entropia (H) com capacidade do canal (C) na expressão 2.6, seria necessário a determinação desta capacidade de canal (C), pois já havia uma fórmula para o cálculo da entropia (H) (equação 2.1). Juntando as "peças" deste "quebra-cabeças", Shannon propôs a equação 2.7 como método de cálculo da capacidade de canal (SHANNON, 1948, p. 43):

$$C = W \log_2 \left(1 + \frac{P}{N}\right) \tag{2.7}$$

Onde:

C = Capacidade do Canal em *bits*/segundo

W = Largura de Banda em *Hertz*

$\frac{P}{N}$ = Relação sinal-ruído P e N, representado em *Watt*, considerando P como a potência média do sinal e N a potência média do ruído.

Desta forma, "[...] Shannon mostrou que cada canal possui uma taxa máxima para transmitir dados eletrônicos de forma confiável, que ele chamou de capacidade do canal" [22] (GUIZZO, 2003, p. 8, tradução nossa). Este limite tornou-se uma referência fundamental para clarear o trabalho de engenheiros de comunicação apontando a eles as necessidades de conhecer e respeitar os limites físicos de transmissão dos sistemas que estavam desenvolvendo.

Apesar da capacidade de canal não constar na literatura científica propriamente dita como teorema de Shannon[23], ela é pré-requisito para a aplicação do Teorema de Codificação de Canal. A determinação da capacidade de canal é muito significativa para a escolha da codificação adequada e a codificação é a aplicação mais significativa da Teoria da Informação (REZA, 1961; GAPPMAIR, 1999). A codificação aplicada tanto sobre o canal quanto sobre a fonte são propostos por Shannon (1948) como ferramentas de otimização para alcançar maior eficiência em um processo de comunicação. Tais codificações serão abordadas nas seções seguintes desta dissertação.

[22] *"[...] Shannon showed that every channel has a maximum rate for transmitting electronic data reliably, which he called the channel capacity."*

[23] Para uma discussão mais aprofundada sobre essa fórmula ver Rioul e Magossi (2014)

2.5.2 Teorema de Codificação de Fonte

Um importante teorema proposto por Shannon foi o Teorema da Codificação de Fonte, que demonstra a possibilidade de calcular o número de *bits* necessários para descrever um dado de forma única para qualquer fonte (GAPPMAIR, 1999). Shannon propõe através dele que, para transmitir mais informação, não é necessário aumentar a capacidade do canal, mas sim trabalhar numa codificação apropriada na fonte. Desta forma, Shannon demonstrou que seria possível "[...] esmagar as mensagens - economizando assim o tempo de transmissão [...]" [24] (GUIZZO, 2003, p. 9, tradução nossa), através de códigos que realizem compressão de dados. Esta é a proposta de eficiência na comunicação trazida pela codificação (MASSEY, 1984; REZA, 1961), neste caso da fonte de dados. Desta forma, apesar de não fornecer os algoritmos, foi a partir do Teorema de Codificação de Fonte que Shannon apresentou ideias iniciais que fundamentaram o surgimento de códigos para compressão de dados e para detecção e correção de erros (GAPPMAIR, 1999).

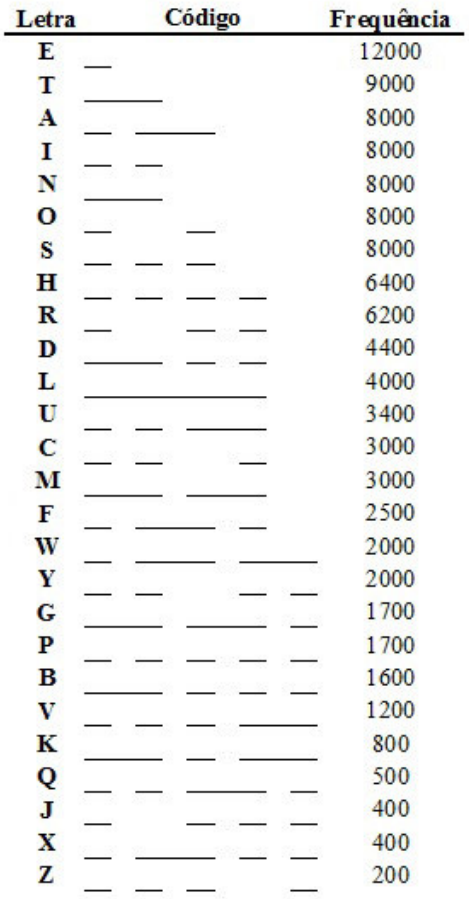
Uma das maneiras mais simples e antigas de comprimir dados na fonte transmissora é realizada através da atribuição de símbolos mais simples (pequenos/curtos) aos resultados mais frequentes e, consequentemente, os maiores/mais longos aos resultados menos frequentes. Um exemplo da aplicação desta técnica é encontrado no código de Samuel Morse (1832)[25]. Neste código, o símbolo mais frequente é representado por um único ponto (COVER; THOMAS, 2012; CHERRY, 1951; BROOKES, 1956; GUIZZO, 2003). A Figura 2.10, é uma adaptação de Cherry (1951, p. 384), em que a coluna da esquerda representa o símbolo, a coluna central representa a codificação e a coluna da direita a relação com as quantidades de cada símbolo encontradas por Morse na impressora de um escritório. Ainda, de acordo com Cherry (1951), esta figura representa o código Morse original.

Inspirado nesta técnica, mas com uma visão mais ampla, Shannon propôs a atribuição de códigos com diferentes comprimentos não apenas para as letras individualmente, mas para pares ou grupos de letras. Conforme disposto na Figura 2.6, considerando a língua portuguesa, a probabilidade de ocorrer a letra U após uma letra Q é muito maior do que qualquer

[24] *"Shannon showed that with specially designed codes engineers could do two things: first, they could squish the messages-thus saving transmission time; also, they could protect data from noise and achieve virtually error-free communication using the whole capacity of a channel-perfect communication at full speed, something no communication specialist had ever dreamed possible"*

[25] O código Morse tornou-se um padrão de comunicação internacional, atualmente recomendado pela *Internacional Telecomumnication Union (ITU-R)*, recomendação nº M.1677-1. Pode ser acessado em: <https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.1677-1-200910-I!!PDF-E.pdf>

Figura 2.10 - Representação do código Morse



Letra	Código	Frequência
E	●	12000
T	▬	9000
A	● ▬	8000
I	● ●	8000
N	▬ ●	8000
O	▬ ▬	8000
S	● ● ●	8000
H	● ● ● ●	6400
R	● ▬ ●	6200
D	▬ ● ●	4400
L	● ▬ ● ●	4000
U	● ● ▬	3400
C	▬ ● ▬ ●	3000
M	▬ ▬	3000
F	● ● ▬ ●	2500
W	● ▬ ▬	2000
Y	▬ ● ▬ ▬	2000
G	▬ ▬ ●	1700
P	● ▬ ▬ ▬ ●	1700
B	▬ ● ● ●	1600
V	● ● ● ▬	1200
K	▬ ● ▬	800
Q	▬ ▬ ● ▬	500
J	● ▬ ▬ ▬	400
X	▬ ● ● ▬	400
Z	▬ ▬ ● ●	200

Fonte: Adaptada de Cherry (1951, p. 384).

outra letra e a probabilidade de ocorrer uma repetição da letra Q é praticamente inexistente. O mesmo pode ser aplicado às palavras em uma frase (GUIZZO, 2003).

Outra observação importante feita por Shannon é a característica estatística dos idiomas, na qual padrões específicos se repetem naturalmente. Tais padrões de repetição podem ser observados, por exemplo, em palavras cruzadas e são previsíveis. A estes padrões Shannon chamou-os de "redundância". No idioma inglês, por exemplo, Shannon indicou haver uma redundância de 50%. Desta forma, ao eliminar a redundância das mensagens, é possível realizar compressão (SHANNON, 1948). Shannon também escreveu em um artigo para a *Encyclopaedia Britannica*, que muitas letras podem ser excluídas sem que seja impossível ao leitor determinar o significado original. Um exemplo interessante é a mensagem " NF RM T N TH R ", a qual induz ao entendimento de "*INFORMATION THEORY*" (GUIZZO, 2003).

Desta forma, ao mapear um alfabeto como o inglês, é importante considerar que a letra “e” aparece repetitivamente muito mais que a letra “q”. Portanto, ao codificar em *bits* este alfabeto, pode-se usar os códigos menores para os que mais aparecem e os maiores para os que menos aparecem. Assim, diminui-se a redundância e realiza-se uma codificação na fonte de dados a fim de comprimir e otimizar a informação antes de ser transmitida (GALLAGER, 2001; COVER; THOMAS, 2012; BROOKES, 1956).

A relação da entropia com compressão e a entropia como pré-requisito para a codificação de fonte é apontada por Brookes (1956) através do seguinte exemplo: Considere-se que uma mensagem aleatória M é formada usando-se o conjunto de símbolos $\{A,B,C,D,E,F,G,H\}$ cujas respectivas probabilidades de ocorrência de cada símbolo na mensagem sejam dadas pelo conjunto $P(M) = \frac{1}{2}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}$. Observa-se que a probabilidade de ocorrer o símbolo A é muito maior que os demais. As Tabelas 2.5 e 2.6 representam duas formas de codificação para essa mensagem, respectivamente chamados de código 1 e código 2. O código 1 utiliza um tamanho fixo de 3 *bits* para representar cada símbolo, enquanto que o código 2 utiliza o número mínimo de *bits* para cada símbolo, de acordo com sua probabilidade de ocorrência. Isso é feito de acordo com uma técnica de compressão, análoga ao código Morse, em que os símbolos que ocorrem como maior probabilidade são representados por uma quantidade menor de *bits*.

Tabela 2.5 - A entropia na compressão - Código 1

Símbolo	Codificação
A	000
B	001
C	010
D	011
E	100
F	101
G	110
H	111

Fonte: (BROOKES, 1956, p. 173)

No primeiro caso, código 1, o tamanho, entropia, de cada símbolo é fixo: $\log_2(8) = 3$ *bits*. O código 2 é melhor que o código 1, pois a entropia é calculada considerando, além do tamanho de cada símbolo, sua probabilidade de ocorrência. Neste caso, e entropia é dada por: $H(M) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{8}\log\frac{1}{8} - 6\cdot\frac{1}{16}\log\frac{1}{16} = 2\frac{3}{8}$ *bits* por símbolo. Desta forma, visualiza-se como

Tabela 2.6 - A entropia na compressão - Código 2

Símbolo	Codificação
A	0
B	110
C	1010
D	1011
E	1000
F	1001
G	1110
H	1111

Fonte: (BROOKES, 1956, p. 173)

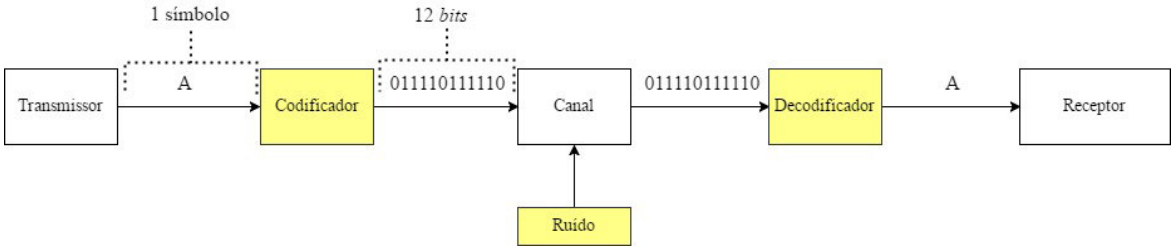
a entropia é usada para medir a eficiência da compressão na codificação de fonte (BROOKES, 1956).

Desta forma, o limite para a compressão de dados é a própria entropia (COVER; THOMAS, 2012). Os exemplos didáticos a seguir demonstram, baseado em Reza (1961), Cover e Thomas (2012), como é realizada uma compressão de dados e também como a entropia é proposta como limite para a compressão. Os exemplos são acompanhados das Figuras 2.11, 2.12 e 2.13 que apresentam visualmente cada exemplo. Estas figuras foram produzidas com base no modelo de comunicação de Shannon, citado por Reza (1961). Nesta representação, os efeitos do ruído são desprezados, pois os códigos de verificação e correção de erros serão tratados mais adiante e neste momento trariam uma complexidade maior e comprometeriam a didática dos exemplos sobre codificação para compressão.

Considera-se S uma *string* binária de valor 011110111110 e tamanho de 12 *bits*. Ela pode ser usada para descrever apenas um símbolo, conforme pode ser observado na Figura 2.11. No entanto, na Figura 2.12 é possível perceber que quando utiliza-se uma codificação melhor, com o mesmo tamanho de 12 *bits*, torna-se possível descrever até seis símbolos diferentes. No terceiro exemplo é aplicada a redundância. Neste caso, a Figura 2.13 mostra a mesma *string* utilizada para descrever símbolos repetitivos, podendo representar até 12 símbolos, ou seja, o limite passa a ser o próprio tamanho. Portanto, quando considerado o tamanho da *string* S como sua entropia, $H(S) = 12$ bits, a codificação máxima de fonte alcançada seria limitada à entropia.

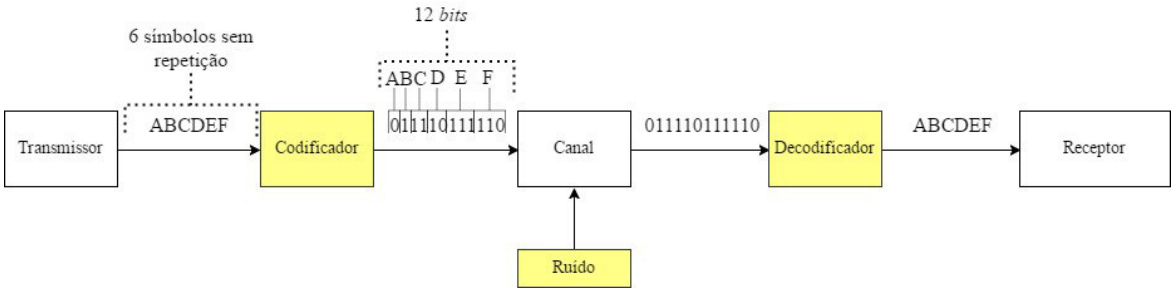
Neste exemplo, quando cada *bit* representa um símbolo, Cover e Thomas (2012) menciona que as atribuições mínimas foram encontradas[26].

Figura 2.11 - Exemplo didático de compressão de dados: 12 *bits* para 1 símbolo



Fonte: Produção do próprio autor.

Figura 2.12 - Exemplo didático de compressão de dados: 12 *bits* para 6 símbolos



Fonte: Produção do próprio autor.

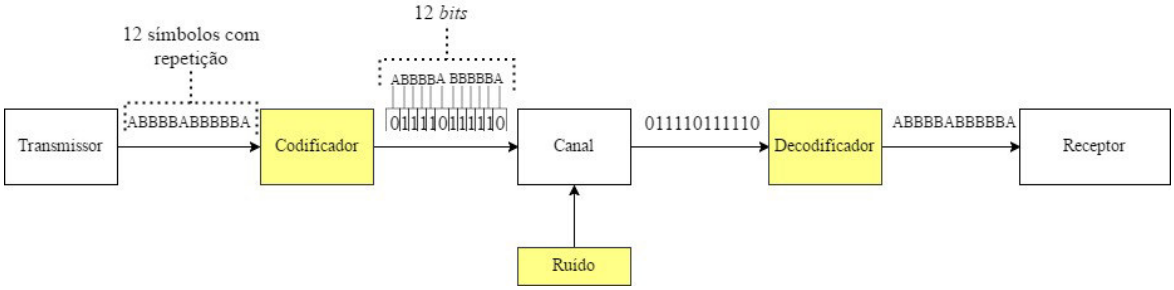
Assim sendo, a entropia é estabelecida como uma medida "natural" de comprimento da descrição mais eficiente, ou seja, a codificação mais eficiente para compressão de dados é aquela na qual cada símbolo pode ser representado por apenas 1 *bit*, o que Shannon chamou de Entropia Máxima.

Quando a entropia máxima não é alcançada, obtêm-se então a Entropia Relativa. Nestes casos, de acordo com Brookes (1956) e Pineda (2006), a eficiência da codificação pode ser calculada pela divisão da entropia relativa pela máxima. Por isso, a redundância, ou seja, o que pode ser removido da mensagem que não prejudicará seu conteúdo, é a medida que complementa a entropia da mensagem até que a entropia máxima seja atingida.

O primeiro código para compressão de dados foi proposto pelo próprio Shannon na seção 9 de seu artigo, porém não foi considerado um código muito eficiente. Logo no ano

[26] Alguns cálculos não são explicitados neste texto, porém há indicativos de como fazê-los em (COVER; THOMAS, 2012) e (REZA, 1961).

Figura 2.13 - Exemplo didático de compressão de dados: 12 *bits* para 12 símbolos



Fonte: Produção do próprio autor.

seguinte, o professor do MIT, Dr. Robert M. Fano, colega de trabalho de Shannon, na busca por códigos eficientes, publica o conhecido código de Fano. No entanto, seu código ainda não era tão eficiente quanto se desejava. Então, o Prof. Fano levou a questão aos alunos de sua turma de teoria da informação. Um de seus alunos, David A. Huffman, motivou-se a encontrar uma solução. Huffman estava a ponto de desistir do desafio, quando teve a ideia de começar a construir uma árvore binária invertida, das folhas para a raiz. Desta forma obteve sucesso e rapidamente provou que seu código era o mais eficiente (MOSER; CHEN, 2012). Atualmente, este código é representado através da construção de um algoritmo para encontrar as atribuições mínimas de comprimento para descrições esperadas, otimizando assim a codificação de dados na fonte (REZA, 1961; COVER; THOMAS, 2012).

Atualmente existem vários códigos para compressão de dados, dentre eles o algoritmo de Huffman (1952) é classificado como um método estatístico e também é o mais utilizado, pois é encontrado na maioria dos programas de computadores pessoais. A construção deste se dá através de uma árvore binária em que os galhos mais próximos da raiz são ocupados por símbolos que possuem uma frequência maior de ocorrência. Consequentemente, os símbolos mais frequentes serão representados por *strings* binárias mais curtas enquanto os símbolos menos frequentes pelas *strings* mais longas, ou seja ao montar a codificação de um conjunto de símbolos em uma *string* binária, os símbolos que mais se repetem utilizarão a menor quantidade de *bits* (SALOMON, 2012). Desta forma, a proposta de Huffman (1952) é fundamentada na teoria de Shannon em atingir o mínimo de redundância possível na fonte.

O próximo capítulo desta dissertação exibe exemplos didáticos e práticos que permitem compreender a aplicação dos conceitos da codificação da fonte de dados aplicados às tecnologias atuais de informação e comunicação.

2.5.3 Teorema da Codificação de Canal

Contradizendo a proposta da eliminação da redundância na codificação de fonte, a codificação de canal entende que a redundância é necessária e nem sempre indesejável. Na codificação de canal a proposta é que a redundância seja inserida na mensagem como forma de proteger seu conteúdo de possíveis erros. Inicialmente pode parecer estranho mas, foi a proposta de Shannon para enfrentar os erros na transmissão através da codificação para detecção e correção de erros (GUIZZO, 2003).

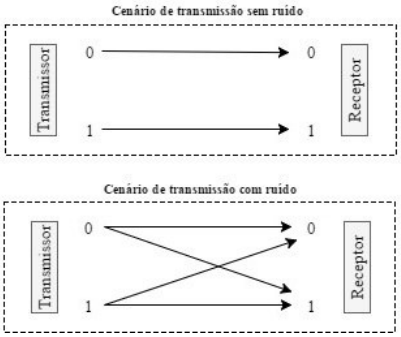
Atualmente, o uso de códigos para detecção e correção de erros é muito comum não só para transmissão, mas também para armazenamento de informações digitais, existindo diferentes tipos de códigos que realizam esse processo (LIN; COSTELLO, 1983). Em ambos os contextos, tanto de transmissão quanto de armazenamento de informações, a importância da utilização destes códigos se dá pela característica dos erros serem espontâneos e inevitáveis.

No contexto de armazenamento, os *chips* de memórias de computadores são construídos em silício de modo que os impulsos sejam representados por 0 ou 1, através da retenção ou ausência de voltagem elétrica. Nesta construção, a grande vilã é a partícula Alfa: um núcleo radioativo encontrado em pequena quantidade em quase todos os materiais. Tal partícula é encontrada tanto nos *chips* de memórias como também em outros materiais como na própria embalagem plástica deste *chip*. Desta forma, as memórias estão constantemente submetidas a este bombardeio atômico que pode afetar o movimento dos elétrons, fazendo com que os 1´s armazenados tornem-se 0´s ou vice-versa, consequentemente causando erros espontâneos e inevitáveis. Assim, informações digitais armazenadas nestas memórias sem um mecanismo de detecção e correção de erros perderiam a confiabilidade (MCELIECE, 1985).

No segundo contexto, das transmissões, os fenômenos ocorridos durante o processo que tornam a mensagem imprevisível são chamados de ruídos (FANO, 1950). Em sistemas de comunicação, os ruídos são impossíveis de serem totalmente eliminados, por isso são considerados como um de seus principais “vilões”. Tais ruídos são causados devido a interferências de fios adjacentes em telefonia, tempestades com trovoadas, tempestades magnéticas, correntes espúrias dentro de equipamentos, etc. Até mesmo em canais ópticos com fibra óptica, as perdas de energia que degradam a luz transmitida causam ruídos (GUIZZO, 2003). A Figura 2.14 ilustra a dificuldade de transmissão digital associada ao ruído.

Assim como no exemplo do Código Morse, citado na seção anterior como procedimento antigo de codificação para compressão de dados, a codificação para detecção e correção

Figura 2.14 - Cenários de transmissão com e sem ruído



Fonte: Adaptada de Cover e Thomas (2012, p. 185).

de erros também tem exemplos muito antigos. Dentre eles, um dos mecanismos mais óbvios usados antes do artigo de Shannon era a repetição de blocos de símbolos com a mesma informação: a redundância (COVER; THOMAS, 2012).

> Naquela época, a maneira usual de superar o ruído era aumentar a energia dos sinais de transmissão ou enviar a mesma mensagem repetidamente [...] quando Shannon apresentou uma maneira muito mais eficaz para evitar erros sem o desperdício de tanta energia e tempo: a codificação. [27] (GUIZZO, 2003, p. 9, tradução nossa)

Desta forma, o Teorema da Codificação de Canal propõe uma maneira eficiente de reduzir a taxa de erro em uma transmissão por um canal ruidoso (GAPPMAIR, 1999; PIERCE, 1973), ao invés das práticas utilizadas até o momento que sobrecarregavam os canais e dificultavam as transmissões, além de aumentar o custo, exigir mais energia, o que refletia no tamanho das baterias ou outras fontes de eletricidade (GUIZZO, 2003).

No Teorema de Codificação de Canal o problema fundamental da comunicação de "[...] reproduzir em um ponto exatamente ou aproximadamente uma mensagem selecionada em outro ponto [...]" [28] (SHANNON, 1948, p. 1, tradução nossa), enfrentando os erros, também é resolvido através da codificação. Neste caso, o "[...] objetivo da codificação é introduzir redundância de modo que, mesmo que algumas das informações sejam perdidas ou corrompidas, ainda será possível recuperar a mensagem no receptor" [29] (COVER; THOMAS, 2012, p. 210,

[27] *"At that time, the usual way to overcome noise was to increase the energy of the transmission signals or send the same message repeatedly-much as when, in a crowded pub, you have to shout for a beer several times. Shannon showed a better way to avoid errors without wasting so much energy and time: coding."*

[28] *"[...] fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point."*

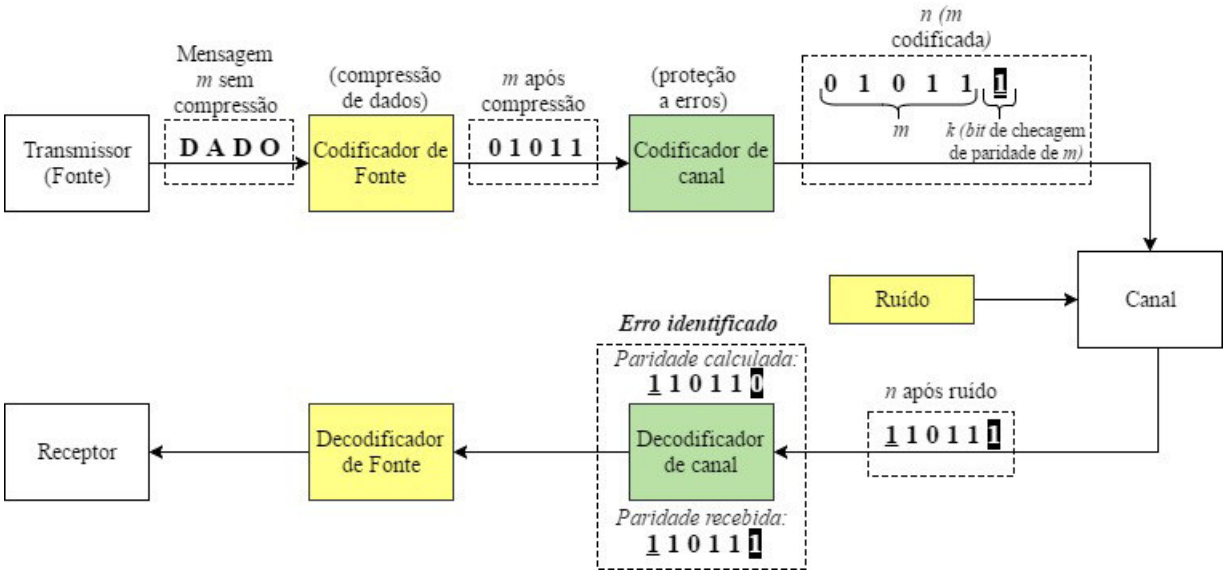
[29] *"The object of coding is to introduce redundancy so that even if some of the information is lost or corrupted, it will still be possible to recover the message at the receiver."*

tradução nossa). No entanto, essa redundância é inserida na mensagem de maneira estratégica e assim como na codificação de canal, a entropia é fundamental para esta implementação.

Desta forma, a entropia não é mais usada para medir toda a mensagem, mas sim realizar medidas parciais, de posições específicas da mensagem. Tais medidas obedecem regras de distribuição de probabilidades que possibilitam a detecção de erros ao realizar novamente estas medidas no receptor (BROOKES, 1956). A forma mais simples de realizar essas medições parciais, de acordo com Hamming (1950), é através da checagem de paridade.

O processo de checagem de paridade consiste em distribuir os dígitos de uma mensagem de tamanho m em um vetor de n dígitos, onde a n-ésima posição seja um dígito específico adicional para esta checagem. Como o próprio nome sugere, a checagem de paridade verifica se a soma de dígitos 1's da mensagem é par. Caso esta verificação seja verdadeira, atribui-se 0 ao dígito de paridade, caso contrário atribui-se 1. Por exemplo, a mensagem $m = 0101$ poderia ser codificada em $n = 01010$, onde o último dígito é 0, pois a quantidade de dígitos 1 da mensagem é par. Em um processo de transmissão, em que um ruído altere algum dígito desta mensagem de 0 pra 1 ou vice-versa, a checagem de paridade seria capaz de detectar o erro. A Figura 2.15 ilustra o exemplo utilizado na seção anterior, a palavra "DADO" comprimida com o algoritmo de Huffman (1952), sendo codificada com dígito de checagem de paridade.

Figura 2.15 - Exemplo de checagem de paridade



Fonte: Adaptada de Lin e Costello (1983, p. 2).

No entanto, a checagem de paridade aplicada como no exemplo citado na Figura 2.15 torna-se ineficiente na ocorrência de mais de um erro. Também é insuficiente para indicar exatamente em qual *bit* o erro ocorreu, inviabilizando assim o processo de correção. Desta forma, Hamming (1950) propõe que a paridade não seja checada em todos os dígitos, mas somente em posições selecionadas.

Considerando que Shannon não apresentou os métodos e algoritmos em seu artigo de 1948, nos anos que se seguiram eles foram brilhantemente desenvolvidos por outros pesquisadores (MOSER; CHEN, 2012). Esse desenvolvimento buscava um código bom e simples, necessitando ser bom de modo a garantir baixas probabilidades de erro e simples para permitir a codificação e decodificação facilmente (COVER; THOMAS, 2012). Assim como Huffman apresentou um modelo de codificação para compressão de dados na fonte, Hamming apresentou a primeira proposta de código completo para verificação e correção de erros, considerado um dos mais simples métodos de transmissão na presença de ruído (REZA, 1961). Atualmente existem diferentes tipos de códigos que realizam esse processo. Os mais comuns são os códigos de bloco e os códigos convolucionais. O código de Hamming é classificado como código de bloco (LIN; COSTELLO, 1983) e foi escolhido nesta dissertação para indicar, no próximo capítulo, os impactos da Teoria da Informação nas TICs.

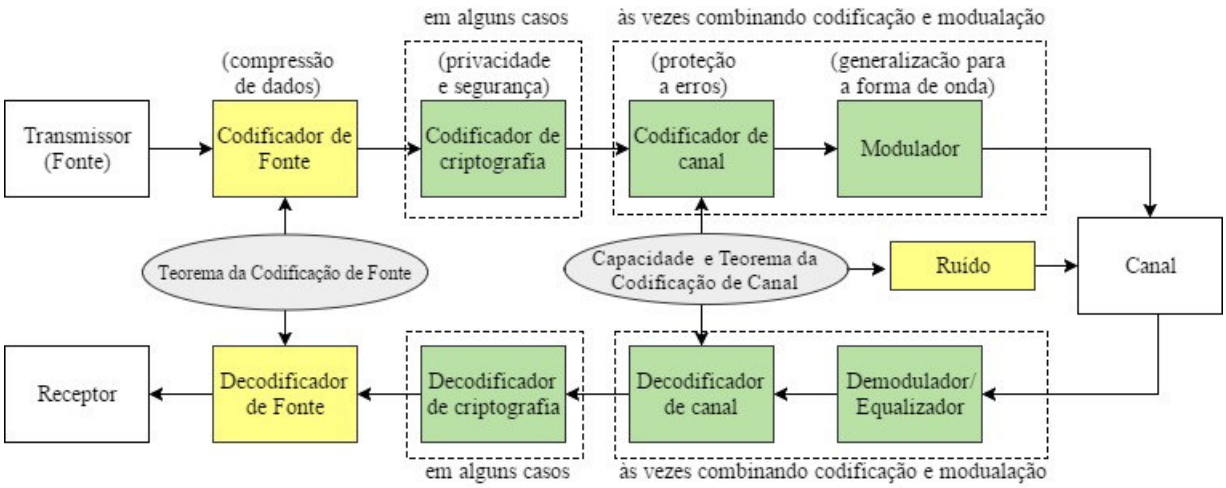
Ao longo deste capítulo foram acrescentados novos elementos ao modelo "natural" de comunicação expresso através de um diagrama de blocos na Figura 2.2. Inicialmente, na Figura 2.3 acrescentou-se elementos associados a conceitos gerais trazidos da Teoria da Informação. Logo a seguir, nas Figuras 2.11, 2.12 e 2.13 novos elementos associados a codificação de fonte (compressão) foram adicionados. Agora, a complexidade do modelo precisa ser ampliada ainda mais através do acréscimo dos elementos que fazem a codificação de canal. Esse novo modelo é ilustrado na Figura 2.16.

A codificação de canal para detecção e correção de erros pode ser observada, de maneira didática, no *site*[30] do Departamento de Matemática Aplicada e Informática da Universidade Técnica da Dinamarca. Neste *site* é disponibilizado um simulador de codificação e decodificação de canal.

Como primeiro passo neste simulador, escolhe-se uma foto ou até mesmo é possível enviar uma foto para ser codificada. No segundo passo, escolhe-se a velocidade de codificação.

[30]<http://www2.mat.dtu.dk/people/T.Hoeholdt/DVD/index.html>

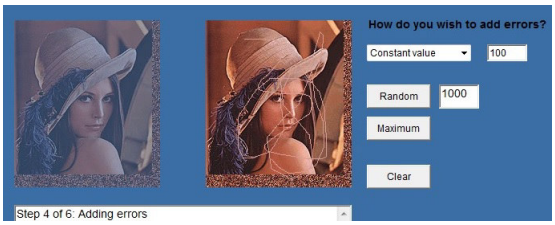
Figura 2.16 - Modelo de um sistema de comunicação com codificações



Fonte: Adaptada de Moon (2005, p. 6) e Lin e Costello (1983, p. 2).

A seguir, com a foto já codificada, pode-se produzir erros através de rabiscos utilizando o *mouse*, conforme observa-se na Figura 2.17.

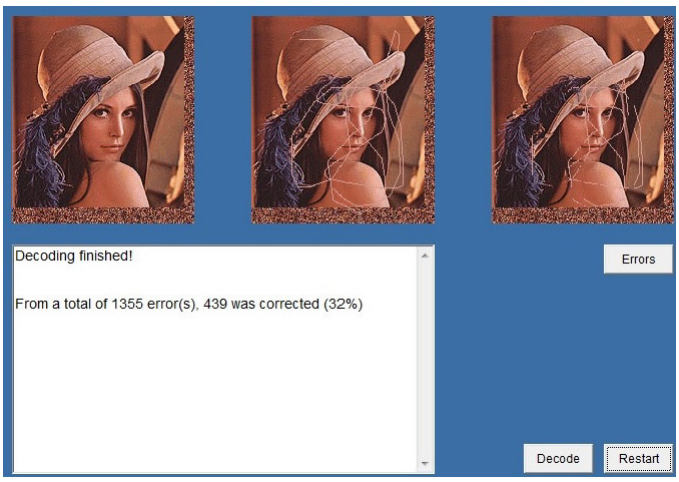
Figura 2.17 - Simulador de codificação de canal: Adicionando-se erros (rabiscos)



Fonte: Extraída de DTU (sem data).

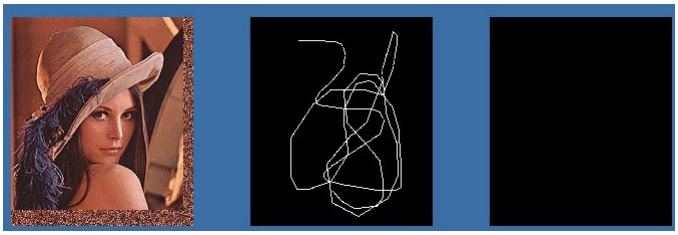
Ao avançar, para a penúltima etapa, é possível escolher um dos dois algoritmos de detecção e correção de erros disponibilizados no *site* e aplicar as correções através do processo de decodificação. O botão *"Decode"* pode ser usado repetitivamente até que todos os erros sejam eliminados, conforme ilustrado na Figura 2.18. Também é disponibilizado o botão "*Errors*", que exibe as alterações realizadas, em destaque sob fundo preto, como exposto na Figura 2.19.

Figura 2.18 - Simulador de codificação de canal: Resultado da decodificação



Fonte: Extraída de DTU (sem data).

Figura 2.19 - Simulador de codificação de canal: Erros detectados destacados



Fonte: Extraída de DTU (sem data).

2.6 A relação da Teoria da Informação com áreas diversas

De acordo com Eco (2001), a Teoria da Informação se vale de princípios comuns a outras disciplinas da ciência. Tanto o conceito de informação quanto o conceito de comunicação, objetos da Teoria de Shannon, possuem uma grande amplitude, fazendo-se presentes em diversas áreas de conhecimento e campos científicos. Por isso, segundo Verdu (1998) a Teoria da Informação pode ser considerada como uma teoria unificadora com intersecções profundas com probabilidade, estatística, ciência da computação entre outras. Além dessas, Brookes (1956) cita outras áreas que usam conceitos de Teoria da Informação como: neurologia e biblioteconomia, mecânica estatística, psicologia, criptografia e biologia.

O interesse de outras áreas pela Teoria da Informação fora percebido por Shannon logo após a publicação de seu artigo de 1948. Ele recebeu cartas de vários locais de seu país e do mundo como Canadá, Inglaterra, França e Japão. Com perguntas, ou simplesmente

ideias compartilhadas, vindas de famosas universidades, laboratórios de empresas, instituições governamentais e militares. Com o passar do tempo, os remetentes destas correspondências não seriam apenas matemáticos ou engenheiros, mas seriam também economistas, psicólogos, sociólogos e linguistas (GUIZZO, 2003).

Dentre as diversas relações da Teoria da Informação com outras áreas, destacam-se:

1. Engenharia Elétrica (Teoria da Comunicação): Melhorias na capacidade de transmissão de um canal para que a capacidade possa ser calculada a partir das características de ruído e com probabilidades de erros insignificantes (COVER; THOMAS, 2012).

2. Ciência da Computação (Complexidade de Kolmogorov): traz a ideia de que a complexidade de uma série de dados pode ser definida pelo comprimento do programa de computador binário mais curto para computar a *string*. Assim, a complexidade é o comprimento mínimo da descrição. Desta forma, a complexidade de Kolmogorov é parecida com entropia de Shannon. É uma compressão de dados que leva a um procedimento logicamente consistente para inferência (COVER; THOMAS, 2012).

3. Física (Termodinâmica): A ideia de que a entropia só aumenta. A mecânica estatística como área de nascimento da entropia e da Segunda Lei da Termodinâmica (BRILLOUIN, 1956; COVER; THOMAS, 2012).

4. Matemática (Teoria da Probabilidade e Estatística): As ideias de entropia, entropia relativa e informação mútua são definidas como funcionais das distribuições de probabilidade. Por sua vez, elas caracterizam o comportamento de longas sequências de variáveis aleatórias e nos permitem estimar as probabilidades de eventos raros e encontrar o melhor expoente de erro em testes de hipóteses - *large deviation theory* (COVER; THOMAS, 2012).

5. Filosofia da Ciência (*Occam's Razor*): Traz a ideia de que as causas não devem ser multiplicadas além das necessidades, ou *o que é mais simples é melhor* (COVER; THOMAS, 2012).

6. Economia (Investimento): O investimento repetido em um mercado de ações estacionário resulta em um crescimento exponencial da "riqueza". A taxa de crescimento da riqueza está relacionada à taxa de entropia do mercado de ações. Os paralelos entre a teoria do investimento ideal no mercado de ações e a teoria da informação são impressionantes

(COVER; THOMAS, 2012). Investidores e analistas de mercado utilizam a teoria da informação para estudar o comportamento do mercado de ações e otimizar um portfólio de ações (GUIZZO, 2003). Técnicas baseadas na Teoria da Informação são aplicadas para a avaliação de riscos de investimentos e para a distribuição de capital em ações (TAKADA; SANTOS, 2014; TAKADA, 2016).

7. Biologia: Os geneticistas e biólogos moleculares usam a teoria da informação para estudar o código genético e investigar hipóteses do sexo como estratégia evolutiva e vencedora para muitas espécies (GUIZZO, 2003). Também utilizada para estudar sociedades de insetos (BROOKES, 1956).

8. Linguística: Alguns cientistas estudaram a orelha e os olhos humanos como canais de informação e calcularam suas supostas capacidades em *bits*. A orelha: 10.000 bits por segundo; O olho: 4 milhões de bits por segundo. Uma pessoa falando: 30.000 bits por segundo. Também é usada por linguistas para a determinação do número de palavras de uma linguagem e seus respectivos comprimentos (GUIZZO, 2003). Outra aplicação importante existente na Linguística é a relação entre Semiótica e Teoria da Informação (ECO, 1976).

9. Psicologia: Um grupo de psicólogos mediu o tempo de reação de uma pessoa para várias quantidades de informações (GUIZZO, 2003).

Como observa-se nos exemplos citados, os impactos da Teoria da Informação superam as barreiras do processo de comunicação entre máquinas. O próprio Shannon alguns anos depois reconheceu que as conexões da Teoria da Informação com campos tão elegantes como computação, cibernética e automação, despertaram a publicidade popular e científica (SHANNON, 1956). Consequentemente, em poucos anos, a Teoria da Informação concebida com a intenção de ser uma ferramenta técnica para o engenheiro de comunicação, transformou-se em um movimento científico em que cientistas de vários campos passaram a usar as ideias da Teoria da Informação em seus próprios problemas e em aplicações para biologia, psicologia, linguística, teoria da organização e muitos outros.

Este impacto foi percebido pelo aumento da proliferação de artigos voltados à Teoria da Informação após 1948. Algumas associações e revistas científicas decidiram limitar o escopo de publicações ditas serem da área de Teoria da informação (NEBEKER, 1998b). O assunto Teoria da Informação tornou-se tão vasto que qualquer coisa estava no escopo dessa teoria

e passivo de publicação, desta forma os editores de revistas da área de engenharia passaram por dificuldades e começaram a determinar uma política que restringia o tema dos artigos (AFTAB *et al.*, 2001). É o caso, por exemplo, do IEEE, que publica a revista *IEEE Information Theory Society*[31], dedicada às publicações voltadas, grosso modo, à Teoria de Shannon. Há também a revista *Entropy*[32], que aborda o conceito de entropia e suas inúmeras aplicações. Por isso, ao realizar uma busca nas bases de indexação de periódicos mais conceituados é comum encontrarmos essa diversidade de áreas utilizando conceitos da Teoria da Informação como metodologia para a resolução de problemas específicos, como por exemplo:

1. Transferência de conhecimento (ALBINO; GARAVELLI; GORGOGLIONE, 2004).
2. Biologia: evolução (BEECHER, 1989; MOUILLOT; LEPRETRE, 1999); população (CATTADORI; HAYDON; HUDSON, 2005); ecologia (MOUTON *et al.*, 2008); genética (SHPAK; CHURCHILL, 2000).
3. Telecomunicações e Ciência da Computação (PAK; PAROUBEK, 2010).
4. Economia e sustentabilidade (TEMPLET, 1999).

Os resultados são ainda maiores quando é feita uma busca simples na Internet utilizando a expressão *"Information Theory"*. São exibidos mais de 5 milhões de resultados, montante pelo qual deduz-se a magnitude da Teoria da Informação e seus impactos na humanidade através de diversos seguimentos científicos.

A vantagem desta rápida expansão da Teoria da Informação é que mesmo sendo "[...] aplicada indiscriminadamente a todos os tipos de sistemas de comunicação, incluindo linguagem, estimulou novas formas de pensar sobre o armazenamento e transmissão de todo tipo de informação no sentido mais geral dessa palavra" [33](BROOKES, 1956, p. 170, tradução nossa). No entanto, apesar de considerar essa popularidade agradável e emocionante e acreditar que muitos dos conceitos da Teoria da Informação são muito úteis a outros campos, Shannon (1956) a considera também perigosa, apontando cuidados a serem observados como:

- O uso de algumas palavras populares como informação, entropia, redundância, não resolve todos os problemas.

[31] <http://www.itsoc.org>

[32] <http://www.mdpi.com/journal/entropy>

[33] *"Although it has been applied indiscriminately to all kinds of communication systems including language, it has nevertheless stimulated new ways of thinking about the storage and transmission of every kind of information in the most general sense of that word."*

- O núcleo da Teoria da Informação é um ramo da matemática e portanto uma compreensão completa da base matemática é um pré-requisito para a expansão a outras aplicações.
- Apesar de lento e tedioso o processo de hipótese e verificação experimental, toda nova aplicação deve ser testada em uma grande variedade de situações experimentais.

Capítulo 3

O conceito de entropia presente nas TICs

Conforme abordado no capítulo anterior, é possível notar que "A Teoria Matemática da Comunicação" de Shannon (1948) se faz presente em muitas áreas, independentemente desta presença ocorrer de maneira direta ou indireta. Ocorre de maneira direta quando há utilização de seus conceitos matemáticos ou probabilísticos nas aplicações específicas das áreas diversas, conforme exemplos apresentados na seção anterior. A presença indireta ocorre pela utilização das TICs, pois atualmente com a expansão do conceito de *Internet das Coisas*, pode-se estimar, utilizando parte de uma citação muito conhecida de Boltzmann (1974, p. 20, tradução nossa), que: não é impossível, apenas "[...] infinitamente improvável [...]"[1] que alguma área não use TICs. E, onde há TICs, provavelmente há Teoria da Informação e matemática. Desta forma, direta ou indiretamente a Teoria da Informação faz-se presente e a matemática se mostra como uma linguagem apropriada aos desenvolvimentos tecnológicos.

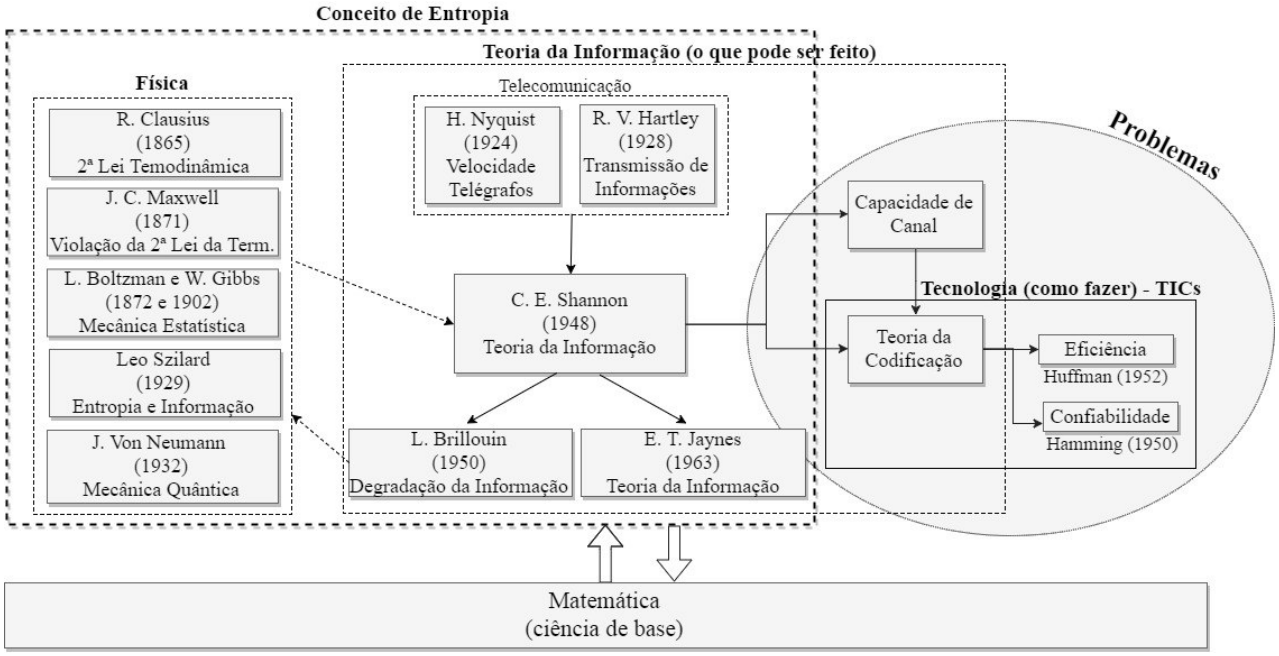
Portanto, ao considerar a Teoria da Informação como exemplo de elo de conexão entre a matemática e a tecnologia, bem como as aplicações tecnológicas da Teoria da Informação estarem alicerçadas no conceito de entropia, a Figura 3.1 ilustra, através de um mapa mental, a proposta de discussões da próxima subseção que visa indicar a presença da Teoria da Informação em setores diversos da sociedade através da codificação de Huffman e Hamming utilizada para resolver problemas de eficiência e confiabilidade no uso das TICs. A seguir abordam-se outros possíveis problemas que envolvam conceitos de entropia e Teoria de Informação, bem como um problema de interesse futuro identificado neste trabalho de pesquisa: *Broadcast Channel* (ou canais de difusão).

3.1 A presença da Teoria da Informação em setores diversos da sociedade

A busca por um equilíbrio entre velocidade de transmissão e taxa de erro demonstrava a necessidade de tecnologias que resolvessem problemas associados à eficiência e confiabilidade das informações. A teoria matemática de Shannon proporcionou tal equilíbrio com a

[1] *"[...] is the most infinitely improbable configuration of energy conceivable [...]"*

Figura 3.1 - Mapa mental das discussões do capítulo 3



Fonte: Produção do próprio autor.

redundância usada de forma mais precisa através de seus teoremas de codificação. Essa precisão parte do princípio de inicialmente remover a quantidade redundante e desnecessária de *bits* de uma mensagem, através de técnicas de compressão. Em seguida, utiliza técnicas para detecção e correção de erros que adiciona o tipo certo de redundância para garantir a confiabilidade da informação (GUIZZO, 2003).

Um dos primeiros exemplos das aplicações práticas da Teoria da Informação, o CD, causou uma grande revolução. Foi um dos primeiros mecanismos de armazenamento de dados que implantou de maneira eficiente algoritmos que fazem tanto a compressão de dados quanto a detecção e correção de erros. Neste caso, a compressão de dados é realizada através da gravação do áudio em formato digital. Desta forma, quando a música é discretizada, convertida de sua fonte natural, contínua, para digital através de um processo de amostragem, são eliminadas frequências inaudíveis aos humanos, o que pode ser considerada uma redundância, algo desnecessário na mensagem para sua compreensão. A detecção e correção de erros é implantada através da inserção de redundâncias necessárias para que as músicas possam ser ouvidas mesmo que hajam pequenos riscos na superfície do disco (GUIZZO, 2003; PINEDA, 2006).

A técnica de detecção e correção de erros utilizada em um CD é a mesma que fora utilizada na espaçonave *Pionner* pela NASA, durante a corrida espacial no período pós-

guerra fria, como técnica implementada que permitia que a espaçonave alcançasse distâncias maiores da Terra e ainda pudesse haver comunicação (GUIZZO, 2003). Apesar de ser um bom exemplo da presença da Teoria da Informação nas TICs, CDs e DVDs são exemplos muito antigos, das décadas de 80 e 90 e que estão em desuso na atualidade. No entanto, é importante o destaque que esse desuso ocorreu pela evolução de outras tecnologias também fundamentadas pela Teoria da Informação como o MP3, o WMA, AVI, MPG, HDTV e as transmissões de áudio e vídeo instantâneas pela Internet, através por exemplo do *Netflix, Youtube* e outros canais de transmissão digital.

3.1.1 Código de Huffman

A ilustração da aplicação do algoritmo de Huffman (1952), bem como as etapas para construção da árvore binária é feita através dos seguintes exemplos didáticos simplificados, em que considera-se cada letra como um símbolo, não fazendo relação neste momento da letra com a Tabela ASC. Tais exemplos também são acompanhados das Figuras 3.2 e 3.3 que representam estas ilustrações. O primeiro exemplo sugere a compressão da palavra "DADO".

Primeiro exemplo - Etapas para a codificação:

1. Frequência da ocorrência: É analisada a frequência de ocorrência de cada símbolo. Em nosso exemplo, o símbolo D é repetido duas vezes, enquanto que os símbolos A e O apenas uma vez.

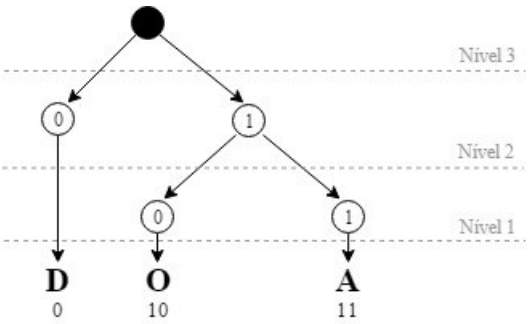
2. Posicionamento nos galhos/nós da árvore binária: Os símbolos com maior frequência são colocados mais próximos da raiz.

3. Codificação dos símbolos: Os caminhos que partem da raíz até o símbolo representam a *string* binária da codificação utilizada.

4. *String* binária codificada: É formada pela junção das *strings* binárias dos símbolos codificados. Na Figura 3.2 observa-se que o símbolo de maior frequência D é representado binariamente por 0, enquanto que os símbolos A e O são representados respectivamente pelas *strings* binárias 10 e 11. Neste caso, a *string* binária utilizada para representar a palavra DADO comprimida poderia ser 0,10,0,11.

Associado à ideia de árvore, o segundo exemplo sugere como mensagem uma palavra que representa o nome de uma árvore: "ARAUCÁRIA". No entanto, para este exemplo

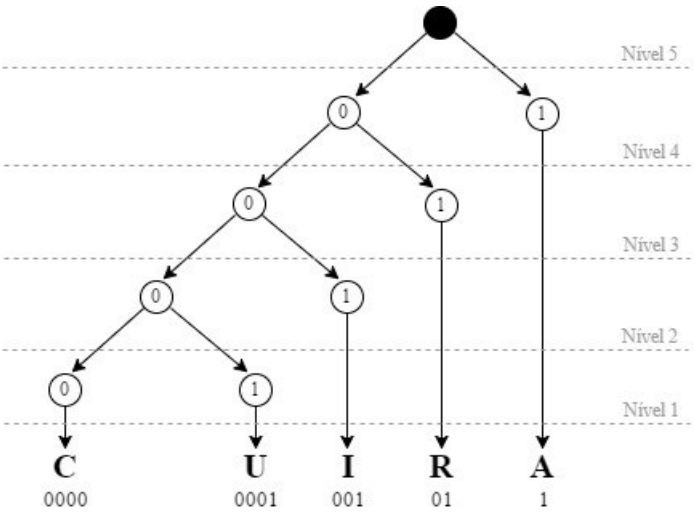
Figura 3.2 - Exemplo didático de compressão com código de Huffman - Palavra “DADO”



Fonte: Produção do próprio autor.

didático o acento agudo da terceira letra “A” será desconsiderado, bem como as representações dos caracteres pela Tabela ASC. A árvore binária da Figura 3.3 ilustra essa compressão.

Figura 3.3 - Exemplo didático de compressão com código de Huffman - “ARAUCARIA”



Fonte: Produção do próprio autor.

As representações acima foram feitas de maneira didática, visando a compreensão mais simplificada do código de Huffman através da construção de uma árvore binária. No entanto, comparando os dois exemplos é dedutivo que quanto maior a expressão que se deseja codificar, maior a árvore binária. Consequentemente, ficará mais complexa a montagem. Por isso, se faz interessante a organização da modelagem matemática proposta por Huffman (1952) em um algoritmo.

Por isso, em seu artigo, Huffman (1952) não propôs explicitamente um algoritmo, mas sim uma modelagem matemática para construção do “*Optimum Binary Code*”. Tal mo-

delagem é baseada nas probabilidades de ocorrência dos símbolos. Considerando N como o número de símbolos de um conjunto, o somatório das probabilidades de ocorrência de cada símbolo (P_i) é igual a 1, ou seja, a certeza é aquela que temos conhecimento de 100% dos símbolos da mensagem. Isso é exposto na expressão 3.1:

$$\sum_{i=1}^{N} P_i = 1 \tag{3.1}$$

Onde:

N = Número de símbolos de uma mensagem

P_i = Probabilidade de ocorrência de cada símbolo na mensagem

Desta forma, Huffman (1952) propõe que o comprimento de um símbolo L_i é o número de dígitos de codificação atribuídos a ele e o comprimento médio da mensagem (L_av) é o somatório da multiplicação das probabilidades de cada símbolo (P_i) pelo comprimento mínimo de cada respectivo símbolo (L_i), conforme a expressão 3.2:

$$L_{av} = \sum_{i=1}^{N} P_i L_i \tag{3.2}$$

Onde:

L_{av} = Comprimento médio da mensagem

N = Número de símbolos de uma mensagem

P_i = Probabilidade de ocorrência de cada símbolo na mensagem

L_i = Comprimento mínimo usado para codificação de cada símbolo

Considerando o modelo matemático de Huffman (1952) no primeiro exemplo didático citado, a probabilidade de ocorrência do símbolo D é de $0,5$ e dos símbolos A e O é de 0,25. Construindo este modelo na Tabela 3.1, os símbolos são dispostos na coluna da esquerda de maneira decrescente em termos de prioridade. A seguir, as probabilidades são somadas a cada nível vertical até o somatório atingir 1, ou seja, a probabilidade que representa a totalidade da mensagem. Na Tabela 3.2 apresentam-se os resultados deste modelo de codificação.

No segundo exemplo, a aplicação do procedimento de otimização binária de Huffman (1952) para a compressão da mensagem "ARAUCARIA" é indicado na Tabela 3.3. Na Tabela 3.4 encontram-se os resultados da aplicação para este segundo exemplo.

Tabela 3.1 - Procedimento de Otimização de Codificação Binária - Exemplo 1

i	Nível1 (P_i)	Nível2	Nível3
D	0,5	**0,5**	**1**
A	**0,25**	**0,5**	
O	**0,25**		

Fonte: Adaptada de Huffman (1952, p. 1100)

Tabela 3.2 - Resultados da Otimização da Codificação Binária - Exemplo 1

i	P_i	L_i	$P_i \cdot L_i$	Código
D	0,5	1	0,5	0
A	0,25	2	0,5	11
O	0,25	2	0,5	10
		$L_{av} =$	1,5	

Fonte: Adaptada de Huffman (1952, p. 1100)

Conforme exposto em Reza (1961), a eficiência E_f da codificação é analisada através da relação entre a entropia $H(x)$ (tamanho da mensagem original) e L_{av} (comprimento médio da mensagem, após a compressão), de acordo com a expressão 3.3. A taxa de compressão é medida por $1 - E_f$.

$$E_f = \frac{H(x)}{L_{av}} \tag{3.3}$$

Onde:

E_f = Eficiência da codificação

$H(x)$ = Entropia

L_{av} = Comprimento médio da mensagem

Desta forma, cabe ressaltar que a entropia, como critério de medida da taxa de informação, é também um requisito para avaliação da eficiência e da taxa de compressão, pois só é possível esta avaliação se houver um método de medição antes e depois da compressão. Por isso, a seguir são apresentados os cálculos que determinam essa otimização para ambos os exemplos citados anteriormente.

Considera-se o primeiro exemplo da palavra “DADO” onde uma mensagem é representada pelo conjunto de símbolos $M_1 = [D, A, D, O]$. Assim, $S_1 = [D, A, O]$ representa o

Tabela 3.3 - Procedimento de Otimização de Codificação Binária - Exemplo 2

i	Nível1 (P_i)	Nível2	Nível3	Nível4	Nível5
A	0,45	0,45	0,45	**0,45**	**1**
R	0,22	0,22	0,22	**0,55**	
U	**0,11**	**0,22**	**0,33**		
C	**0,11**				
I	0,11	**0,11**			

Fonte: Adaptada de Huffman (1952, p. 1100)

Tabela 3.4 - Resultados da Otimização da Codificação Binária- Exemplo 2

i	P_i	L_i	$P_i \cdot L_i$	Código
A	0,45	1	0,45	1
R	0,22	2	0,44	01
U	0,11	4	0,44	0001
C	0,11	4	0,44	0000
I	0,11	3	0,33	001
		$L_{av} =$	2,1	

Fonte: Adaptada de Huffman (1952, p. 1100)

conjunto de símbolos do alfabeto que será utilizado para representar esta mensagem e $P(S_1) = [0,5, 0,25, 0,25]$ representa o conjunto das probabilidades de ocorrência de cada símbolo nesta mensagem. Por fim, $C_1 = [0, 11, 10]$ é o conjunto da codificação binária para representar cada símbolo. Ao aplicar a relação 3.3, obtêm-se:

$$E_{f1} = \frac{H(x)}{L_{av}} = \frac{-(0{,}5 \cdot \log_2(0{,}5)) + (0{,}25 \cdot \log_2(0{,}25)) + (0{,}25 \cdot \log_2(0{,}25))}{(0{,}5 \cdot 1) + (0{,}25 \cdot 2) + (0{,}25 \cdot 2)} = \frac{1{,}5}{1{,}5} = 1 = 100\% \tag{3.4}$$

Agora, considerando o segundo exemplo da palavra “ARAUCARIA”, a mensagem é representada pelo conjunto de símbolos $M_2 = [A, R, A, U, C, A, R, I, A]$ e $S_2 = [A, R, U, C, I]$ é o conjunto de símbolos do alfabeto utilizado para representar esta mensagem. Além disso, $P(S_2) = [0.45, 0.22, 0.11, 0.11, 0.11]$ é o conjunto das probabilidades de ocorrência de cada símbolo nesta mensagem e $C_2 = [1, 01, 0001, 0000, 001]$ é o conjunto da codificação binária para representar cada símbolo. Ao aplicar a relação 3.3, obtêm-se:

$$E_{f2} = \frac{H(x)}{L_{av}} =$$

$$\frac{-(0{,}45 \cdot \log_2(0{,}45)) + (0{,}22 \cdot \log_2(0{,}22)) + (0{,}11 \cdot \log_2(0{,}11)) + (0{,}11 \cdot \log_2(0{,}11)) + (0{,}11 \cdot \log_2(0{,}11))}{(0{,}45 \cdot 1) + (0{,}22 \cdot 2) + (0{,}11 \cdot 4) + (0{,}11 \cdot 4) + (0{,}11 \cdot 3)} = \tag{3.5}$$

$$\frac{2{,}0588}{2{,}1} = 0{,}98 = 98\%$$

Ambos os exemplos são expostos no Apêndice A desta dissertação através de algoritmos desenvolvidos na ferramenta *MATLAB* que possibilitam a comparação "do antes e depois" no processo de codificação de Huffman (1952) para compressão de dados, bem como auxiliam na validação dos cálculos apresentados.

Como mencionado, os exemplos anteriores consideraram cada letra como símbolo, representado por um número de *bits* definido por uma técnica de compressão de dados. Entretanto, os computadores representam esses símbolos em blocos de 8 *bits*, denominados *bytes* ou caracteres. Este mapeamento de símbolos como caracteres para a comunicação por computadores é feito através de um código chamado *American Standard Code for Information Interchange* (ASCII). Por isso, nesta codificação cada símbolo (letra, número, pontuações, controles de texto, espaço etc.) é representado através de uma sequência binária de 8 *bits*. Nesta sequência, 7 *bits* representam a mensagem e 1 *bit* é responsável pela checagem da paridade (MOSER; CHEN, 2012). A checagem de paridade é abordada na próxima seção deste trabalho. A Tabela 3.5 ilustra essa representação dos caracteres em *bytes*.

Outro exemplo, como uma visualização mais prática dos conceitos de compressão, pode ser obtido através de um simples desenho na ferramenta "*Paint*" do Sistema Operacional *Microsoft Windows*, em uma área com dimensão de 2560 x 1920 *pixels*. Esta dimensão é equivalente a uma das qualidades disponíveis para fotografias em *smartphones* e *tablets*, equivalente a 5 *Megapixel*. A Figura 3.4 exibe um comparativo entre esse desenho gravado sob um formato sem compressão (BMP) e sob um formato que utiliza compressão (JPG). No primeiro caso, o tamanho do arquivo ficou em 14 MB (equivalente a 14.336 KB) e, no segundo, o mesmo desenho foi representado com apenas 599 KB, com uma taxa de compressão de 95,82 %.

A taxa de compressão das imagens é variável, no entanto, para mensurar o benefício da compressão, ainda mais próxima ao dia a dia dos usuários de TICs, considera-se que um cartão de memória com capacidade de armazenamento de dados de 2 GB (equivalente a 2.048

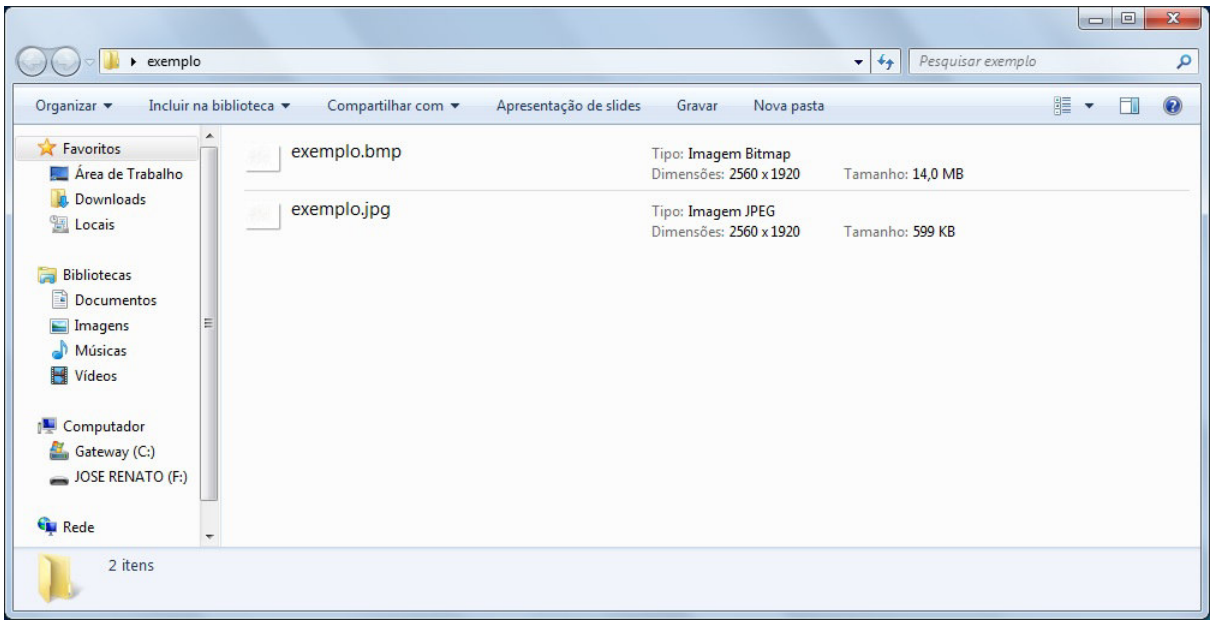
Tabela 3.5 - Tabela ASCII

Binário	Caractere	Binário	Caractere	Binário	Caractere	Binário	Caractere
0000000	NUL	0100000	SP	1000000	@	1100000	‘
0000001	SOH	0100001	!	1000001	A	1100001	a
0000010	STX	0100010	”	1000010	B	1100010	b
0000011	ETX	0100011	#	1000011	C	1100011	c
0000100	EOT	0100100	$	1000100	D	1100100	d
0000101	ENQ	0100101	%	1000101	E	1100101	e
0000110	ACK	0100110	&	1000110	F	1100110	f
0000111	BEL	0100111	’	1000111	G	1100111	g
0001000	BS	0101000	(	1001000	H	1101000	h
0001001	HT	0101001	)	1001001	I	1101001	i
0001010	LF	0101010	*	1001010	J	1101010	j
0001011	VT	0101011	+	1001011	K	1101011	k
0001100	FF	0101100	,	1001100	L	1101100	l
0001101	CR	0101101	−	1001101	M	1101101	m
0001110	SO	0101110	.	1001110	N	1101110	n
0001111	SI	0101111	/	1001111	O	1101111	o
0010000	DLE	0110000	0	1010000	P	1110000	p
0010001	DC1	0110001	1	1010001	Q	1110001	q
0010010	DC2	0110010	2	1010010	R	1110010	r
0010011	DC3	0110011	3	1010011	S	1110011	s
0010100	DC4	0110100	4	1010100	T	1110100	t
0010101	NAK	0110101	5	1010101	U	1110101	u
0010110	SYN	0110110	6	1010110	V	1110110	v
0010111	ETB	0110111	7	1010111	W	1110111	w
0011000	CAN	0111000	8	1011000	X	1111000	x
0011001	EM	0111001	9	1011001	Y	1111001	y
0011010	SUB	0111010	:	1011010	Z	1111010	z
0011011	ESC	0111011	;	1011011	[	1111011	{
0011100	FS	0111100	<	1011100	n	1111100	|
0011101	GS	0111101	=	1011101	]	1111101	}
0011110	RS	0111110	>	1011110	^	1111110	~
0011111	US	0111111	?	1011111	_	1111111	DEL

Fonte: Adaptada de Moser e Chen (2012, p. 20)

MB, ou 2.097.152 KB) seja inserido em um *smartphone* para fotografias digitais. Considera-se também a resolução de 5 *Megapixel* para cada foto, bem como a taxa de compressão alcançada no exemplo da Figura 3.4. Esta capacidade de memória seria suficiente para armazenar aproximadamente 146 fotos em um formato sem compressão. Em um formato que utilize compressão, essa quantidade aumentaria significativamente para 3500 fotos.

Figura 3.4 - Exemplo prático de compressão através do arquivo de um desenho



Fonte: Produção do próprio autor.

Por isso, **ao longo do tempo a codificação para compressão de dados foi incorporada pelas TICs, pois a ideia de Shannon de que a otimização das comunicações se daria pelo código e não pela largura de banda e potência estava correta e a compressão (via o conceito de entropia) permitiu não sobrecarregar os canais no processo de transmissão**. A compressão de dados também é facilmente encontrada nos formatos de arquivos utilizados pelos *softwares* existentes nos computadores, *smartphones*, *smart tvs*, *tablets* e demais equipamentos de TIC, nas extensões: JPG, PNG, MP4, WMV, AVI, FLV, MPEG, MP3, ZIP, RAR entre outras.

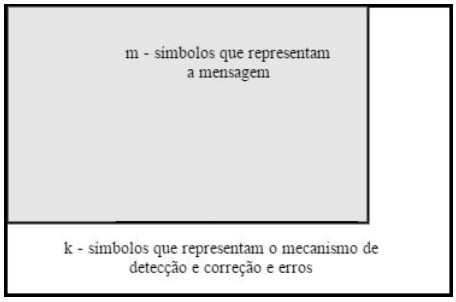
Otimizando ainda mais o processo de comunicação, a compressão de dados também está presente diretamente na transmissão, através da multiplexação dos canais. Uma das famosas técnicas de multiplexação utilizadas, a CDMA, é inspirada na teoria de Shannon (ELLERSICK, 1984). Por isso, a compressão de dados é um exemplo de que as pessoas, mesmo sem terem conhecimento, em suas relações diárias com as TICs, utilizam tecnologias que advém dos conceitos da "Teoria Matemática para Comunicação" de Shannon.

3.1.2 Código de Hamming

Como já apontado anteriormente, os códigos para detecção e correção de erros provocaram uma grande revolução nas telecomunicações, pois possibilitaram, por exemplo, a existência da Internet. Porém, não estão restritos somente às telecomunicações, eles são usados em muitas aplicações comuns como em memórias internas de computadores, celulares, em *pen-drives*, CDs e DVDs. Encontram-se também "disfarçados" sob o nome de dígito verificador, às vezes não associados diretamente à tecnologia, em códigos de barras (UPC), no número do CPF, no ISBN entre outros (MOON, 2005; MOSER; CHEN, 2012).

A proposta de Hamming (1950) utiliza códigos sistemáticos para representar as mensagens que serão transmitidas. Códigos sistemáticos são compostos por símbolos divididos em duas partes: m e k. Desta forma, ao se considerar uma mensagem codificada em n dígitos, a parte que representa a mensagem é m e k são os dígitos utilizados para promover a detecção e correção de erros. A Figura 3.5 representa didaticamente a relação $n = m + k$. Desta relação, surge uma forma para medir a eficiência do código na transmissão: a redundância R. Ela é definida pela relação entre a quantidade de *bits* necessários para a transmissão n e a quantidade de *bits* utilizados para representar a mensagem: $R = \frac{n}{m}$. Ou seja, quanto maior a codificação, maior será a redundância e consequentemente menor o erro.

Figura 3.5 - Exemplo didático de uma foto codificada



Fonte: Produção do próprio autor. Esta figura ilustra didaticamente a aplicação da codificação de Hamming (1950) para detecção e correção de erros em uma foto digital. A parte cinza representa os símbolos que compõem a foto (mensagem), aquela que é visível para as pessoas. A parte branca representa a codificação de detecção e correção de erros

Para Hamming (1950), ao se aumentar o tamanho da mensagem aumenta-se também a complexidade do processo de codificação, de modo que o número k de *bits* necessários para realizar tal codificação pode ser obtido através da inequação 3.6:

$$2^k \geq m + k + 1 \tag{3.6}$$

Onde:

k = número de *bits* necessários para a checagem de paridade

m = tamanho da mensagem em *bits*

Como resultado da aplicação da inequação 3.6, obtém-se a Tabela 3.6.

Tabela 3.6 - Relação quantitativa entre n, m e k

n	m	k
1	0	1
2	0	2
3	1	2
4	1	3
5	2	3
6	3	3
7	4	3
8	4	4
9	5	4
10	6	4
11	7	4
12	8	4
13	9	4
14	10	4
15	11	4
16	11	5
.	.	.
.	.	.
.	.	.

Fonte: Hamming (1950, p. 151)

Além da proporção de k para m, e vice-versa, definidos pela inequação 3.6 e constantes na Tabela 3.6, para que seja possível a identificação precisa do dígito com erro é necessário que os k *bits* utilizados para a checagem de paridade estejam posicionados em n pontos que permitam a distribuição de probabilidades. Tal distribuição ocorre baseada na representação binária da posição do dígito (HAMMING, 1950). Na Tabela 3.7 exibem-se estas distribuições.

Observando-se a Tabela 3.7, Hamming (1950) propõe que os k dígitos necessários para checagem de paridade sejam distribuídos nas posições nas quais há a adição do dígito

Tabela 3.7 - Representação das posições de checagem de paridade e das posições de distribuição dos dígitos *k*

Posição decimal em *n*	Posição binária em *n*	Posições de distribuição dos dígitos *k*
1	00001	2^0
2	00010	2^1
3	00011	
4	00100	2^2
5	00101	
6	00110	
7	00111	
8	01000	2^3
9	01001	
10	01010	
11	01011	
12	01100	
13	01101	
14	01110	
15	01111	
16	10000	2^4
.	.	.
.	.	.
.	.	.

Fonte: Adaptada de Hamming (1950, p. 151-153)

binário 1 à esquerda. Por isso, são distribuídos sequencialmente nas posições em que $2^{log_2(n)}$ é um resultado inteiro, até o limite de 2^k, representados na Tabela 3.7 pelas posições decimais: $1, 2, 4, 8, 16, 32....$

Ainda observando-se a Tabela 3.7, é possível identificar as posições que, de acordo com uma distribuição de probabilidades, devem ser usadas estrategicamente para cada uma das checagens de paridade *k* em *n*. Neste caso, Hamming (1950) propõe que cada checagem seja feita através de equações envolvendo as posições em que o dígito binário seja 1. Por exemplo, a primeira checagem de paridade seria feita considerando a primeira coluna de *bits* da representação binária das posições, considerando somente aquelas em que o valor do *bit* é 1. Portanto, a primeira checagem seria feita nas posições: $1, 3, 5, 7, 9, 11, 13, 15....$ A segunda checagem, seria feita com as posições $2, 3, 6, 7, 10, 11, 14, 15....$ A Tabela 3.8 apresenta um resumo das posições cuja paridade é checada.

As posições de checagem são distribuídas considerando como estratégia as interseções entre elas, permitindo que a detecção de um erro seja realizada através de uma análise

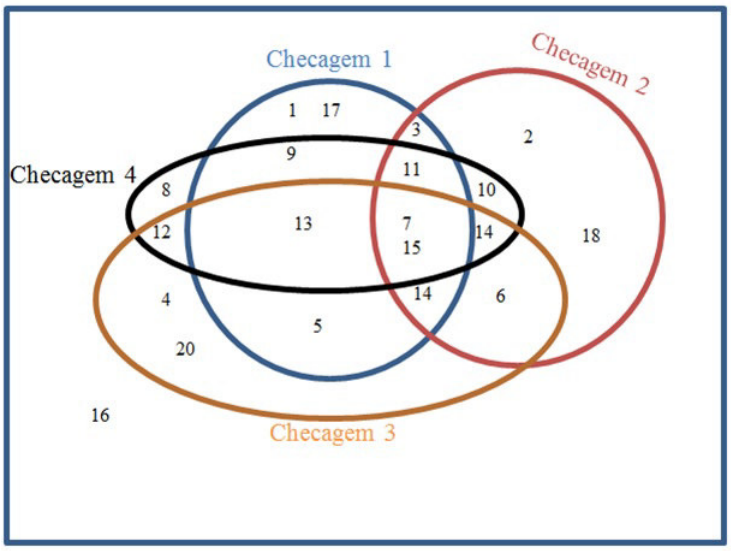
Tabela 3.8 - Representação das posições de checagem de paridade

k	Posição de k em n	Posições de n que são checadas a paridade
1	1	$1,3,5,7,9,11,13,15,17,$ ...
2	2	$2,3,6,7,10,11,14,15,18,$ ...
3	4	$4,5,6,7,12,13,14,15,20,$ ...
4	8	$8,9,10,11,12,13,14,15,24,$ ...
.	.	.
.	.	.
.	.	.

Fonte: Adaptada de Hamming (1950, p. 152)

probabilística que pode ser facilmente notada através de um diagrama de Venn (MCELIECE, 1985; MOSER; CHEN, 2012). Esse diagrama de Venn é exibido na Figura 3.6.

Figura 3.6 - Diagrama de Venn: Distribuição das posições de checagem de paridade



Fonte: Produção do próprio autor.

Portanto, ao se considerar a necessidade de identificar o dígito onde ocorreu o erro para possibilitar a correção, ao se aplicar corretamente as técnicas de codificação de Hamming na *string* binária $m = [01011]$ (mesma utilizada anteriormente na seção sobre codificação de fonte), obtemos a *string* binária codificada $n = [110010111]$. O algoritmo para essa codificação está descrito nas etapas a seguir e exposto no Apêndice B desta dissertação através do *software MATLAB*.

1. Verifica-se o tamanho de m em quantidade de dígitos para se determinar quantos dígitos de checagem de paridade k (redundância) serão necessários para a codificação de m em n. Considerando o exemplo, $m = 5$, de acordo com a Tabela 3.6 são necessários 4 dígitos de checagem de paridade para a codificação de m. Esta relação também pode ser desenvolvida através de cálculo na inequação 3.6.

2. Sabendo que serão necessários 4 dígitos de checagem de paridade, então sabe-se também o tamanho da mensagem codificada n, pois $n = m + k$ e portanto $n = 9$. Os 4 dígitos de checagem, de acordo com as Tabelas 3.7 e 3.8, ocuparão as posições 1, 2, 4 e 8 de n. Portanto, $n = [d_1, d_2, 0, d_3, 1, 0, 1, d_4, 1]$, em que os d_x's representam as posições onde serão inseridos os *bits* de checagem de paridade.

3. Determina-se o valor de cada dígito de paridade d considerando as posições de checagem da Tabela 3.8. Neste caso, considerando p_x as posições dos dígitos de n, as paridades serão determinadas pelas expressões:

   $d_4 = p_8 = 1$ (ímpar) $\rightarrow d_4, p_8 = 1 \rightarrow n = [d_1, d_2, 0, d_3, 1, 0, 1, \mathbf{1}, 1]$

   $d_3 = p_4 + p_5 + p_6 + p_7 = 0 + 1 + 0 + 1 = 2$ (par) $\rightarrow d_3, p_4 = 0 \rightarrow n = [d_1, d_2, 0, \mathbf{0}, 1, 0, 1, 1, 1]$

   $d_2 = p_3 + p_6 + p_7 = 0 + 0 + 1 = 1$ (ímpar) $\rightarrow d_2, p_2 = 1 \rightarrow n = [d_1, \mathbf{1}, 0, 0, 1, 0, 1, 1, 1]$

   $d_1 = p_3 + p_5 + p_7 + p_9 = 0 + 1 + 1 + 1 = 3$ (ímpar) $\rightarrow d_1, p_1 = 1 \rightarrow n = [\mathbf{1}, 1, 0, 0, 1, 0, 1, 1, 1]$

Muito importante também é o processo de decodificação, pois é o responsável por fazer a checagem dos dígitos de paridade k para detectar erros, identificar o dígito errado e corrigi-lo. Como exemplo deste processo considera-se que a mensagem codificada no exemplo anterior $n = [110010111]$ tenha sido afetada por um ruído que transformou o dígito da quinta posição p_5 de 0 para 1. Portanto, n chegou ao decodificador de canal do receptor como $n = [110000111]$.

Neste caso, o processo de decodificação para a detecção e correção de erros, funciona através das etapas do seguinte algoritmo:

1. Verifica-se o tamanho de n em quantidade de dígitos para então determinar a quantidade de dígitos de checagem de paridade k que foram adicionados para a codificação de n. Considerando o exemplo, $n = 9$, a Tabela 3.6 indica a necessidade de 4 dígitos de checagem de paridade.

2. De acordo com a Tabela 3.8, são feitas as checagens de paridade atribuindo-se 0 quando a checagem resultante for verdadeira e 1 quando for falsa. Neste exemplo, com $k = 4$, serão realizadas 4 checagens. As expressões abaixo representam estas checagens, considerando p_x as posições de n utilizadas em cada checagem.

   1ª checagem: $p_1 + p_3 + p_5 + p_7 + p_9 = 1 + 0 + 0 + 1 + 1 = 3$ (ímpar) $\rightarrow \mathbf{1}$

   2ª checagem: $p_2 + p_3 + p_6 + p_7 = 1 + 0 + 0 + 1 = 2$ (par) $\rightarrow \mathbf{0}$

   3ª checagem: $p_4 + p_5 + p_6 + p_7 = 0 + 0 + 0 + 1 = 1$ (ímpar) $\rightarrow \mathbf{1}$

   4ª checagem: $p_8 + p_9 = 1 + 1 = 2$ (par) $\rightarrow \mathbf{0}$

3. O resultado das checagens de paridade, escrito da direita para a esquerda representará, quando convertido de binário para decimal, a posição que contém o erro. Neste exemplo, $0101 = 5$, portanto o erro está na quinta posição. Isto significa que o dígito correto da posição 5 é 1, o que viabiliza o processo de correção.

4. Aplica-se a correção na posição de erro identificada p_5 e são executadas as checagens novamente:

   1ª checagem: $p_1 + p_3 + p_5 + p_7 + p_9 = 1 + 0 + 1 + 1 + 1 = 4$ (par) $\rightarrow \mathbf{0}$

   2ª checagem: $p_2 + p_3 + p_6 + p_7 = 1 + 0 + 0 + 1 = 2$ (par) $\rightarrow \mathbf{0}$

   3ª checagem: $p_4 + p_5 + p_6 + p_7 = 0 + 1 + 0 + 1 = 1$ (par) $\rightarrow \mathbf{0}$

   4ª checagem: $p_8 + p_9 = 1 + 1 = 2$ (par) $\rightarrow \mathbf{0}$

   O resultado das checagens igual a 0000 implica na inexistência de erros.

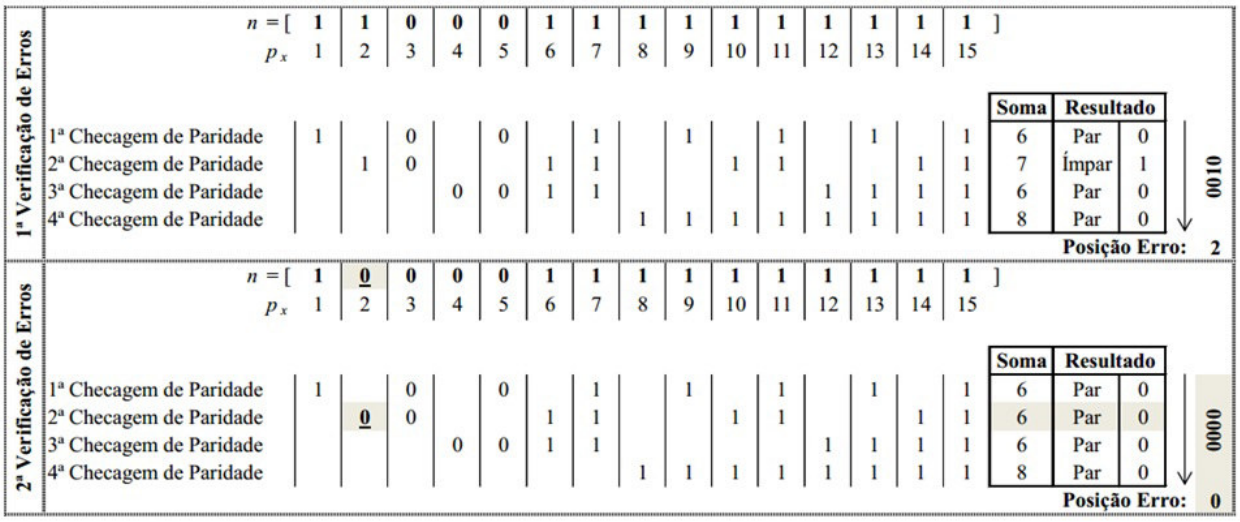
No exemplo anterior, o código correto era previamente conhecido. Agora, em um outro exemplo mais sofisticado (com mais dígitos), supõe-se que foi enviado um código $n = [110001111111111]$. Não se sabe a priori se este código está correto. Mesmo assim, a partir do código de Hamming é possível descobrir se algum desses *bits* está trocado. Aplica-se então as etapas do algoritmo de checagem descritas a seguir e ilustradas pela Figura 3.7.

1. Através da Tabela 3.6, determina-se a quantidade de dígitos utilizados para a correção de erros a partir de n. Neste caso, para $n = 15$, $k = 4$.

2. De acordo com a Tabela 3.8, realizam-se as checagens de paridade atribuindo-se 0 quando a checagem resultante for verdadeira e 1 quando for falsa.

3. O resultado das checagens de paridade é escrito da direita para a esquerda, e quando convertido de binário para decimal indicará o *bit* trocado. Neste exemplo, $0010 = 2$, portanto o erro está na segunda posição, p_2 de n.

Figura 3.7 - Exemplo da aplicação da codificação de Hamming (1950) para a detecção e correção de erros



Fonte: Produção do próprio autor.

Outro exemplo interessante que demonstra a expansão da Teoria da Informação e da Teoria da Codificação, bem como a amplitude do conceito de entropia aplicado a outros campos, é a utilização da codificação de Hamming (1950) para entretenimento. O trabalho de Moreira e Picado (2015) apresenta de forma recreativa alguns truques de magia construídos a partir da codificação, ilustrando de forma didática o alcance e eficácia dos códigos.

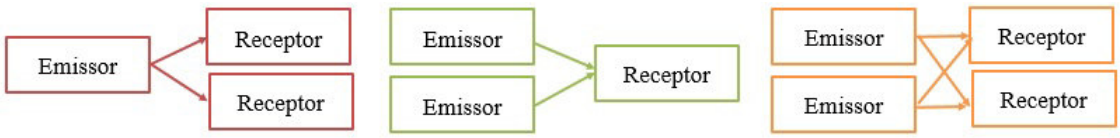
Portanto, o uso da codificação para a detecção e correção de erros, ou para a compressão de dados, são requisitos fundamentais para a existência das tecnologias de informação e comunicação atuais, destacando-se que foi a partir do artigo de Shannon, que grandes esforços foram (e ainda são) realizados para eficiência e controle de erros em um ambiente sujeito a ruídos (LIN; COSTELLO, 1983).

3.2 Problemas abertos que envolvem entropia e Teoria da Informação

Mesmo com a amplitude e plasticidade dos conceitos da Teoria da Informação, principalmente os relacionados à entropia, alguns deles expostos ao longo desta dissertação, ainda há vários problemas em aberto, como pode ser observado, em particular, em (COVER, 1975). Tais problemas estão associados à complexidade da teoria e à sua aplicabilidade em diversos setores onde um modelo de comunicação esteja presente (que se altera de acordo com o desenvolvimentos das TICs). Dentre esses problemas, a determinação da capacidade do canal em transmissões que utilizam múltiplos emissores, receptores e consequentemente canais distintos, proporciona uma quantidade de problemas ainda não resolvidos. Esses problemas impactam nos desenvolvimentos tecnológicos atuais.

Essa nova problemática em que um sistema de comunicação é expandido é chamada de Teoria da informação em Rede (COVER; THOMAS, 2012). Nesta nova óptica o processo de comunicação passa a ter modelos que contém múltiplos transmissores e receptores, conforme pode ser observado na Figura 3.8. Diferentemente do que se apresentou no tradicional diagrama de blocos que consta em (SHANNON, 1948, p. 2), observado em uma adaptação feita por Reza (1961) na Figura 2.3, em que o modelo de comunicação proposto por Shannon, envolve uma única fonte e um único canal.

Figura 3.8 - Comunicações Multicanais



Fonte: Produção do próprio autor.

A inserção destes novos elementos também amplia problemas de comunicação como interferência e ruído. Neste contexto, o problema geral volta a ser o de determinar o "comportamento" da capacidade de canal, bem como as codificações de fonte e de canal nesses tipos de rede. De acordo com Cover e Thomas (2012, p. 509, tradução nossa), "Este problema geral ainda não foi resolvido[...]" [2]. Por isso, pode ser considerado um problema aberto em Teoria da Informação alicerçado no conceito de entropia. Considerando que em 1948, como demonstrado

[2] *"This general problem has not yet been solved[...]"*

no mapa mental da Figura 2.1 e ao longo do capítulo 2, o conceito de entropia possibilitou a determinação da capacidade de canal e consequentemente a codificação para a compressão e para a detecção e correção de erros, estaria também no conceito de entropia a solução para problemas de comunicação em múltiplos canais? Ou seja, reside no conceito de entropia as ferramentas básicas para novos desenvolvimentos em Teoria de Informação em Redes? Essa pergunta não deixa de ser uma proposta para trabalhos futuros e continuação dos assuntos investigados nessa dissertação.

3.2.1 Um problema de interesse: *Broadcast Channel*

Canais de Difusão (*Broadcast Channel*) é um conceito integrante da "Teoria da Informação em Rede". Essa "área" da Teoria da Informação engloba as transmissões em múltiplos canais e classificam-se em: *Multiple-access channel* - Quando existem vários transmissores e apenas um receptor. Um exemplo deste tipo de transmissão seria uma estação de televisão por satélite que transmite sinais de vários pontos a um satélite. *Broadcast* - Significa que existe apenas um transmissor e múltiplos receptores. É o modelo mais comum de transmissões por canais múltiplos (COVER; THOMAS, 2012), encontrado em grande parte dos sistemas de comunicação, o que o torna um importante objeto de estudo.

Cover e Thomas (2012, p. 563, tradução nossa) definem *Broadcast Channel* como "[...] um alfabeto de entrada X e dois alfabetos de saída Y_1 e Y_2 e uma função de transição de probabilidade $p(y_1, y_2|X)$" [3]. Também apresenta o exemplo de um código $((2^{nR_1}, 2^{nR_2}), n)$ para uma transmissão em *Broadcast Channel* onde haveria um codificador $X : (\{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_2}\}) \to X^n$ e dois decodificadores $g_1 : Y_1^n \to \{1, 2, ..., 2^{nR_1}\}$ e $g_2 : Y_2^n \to \{1, 2, ..., 2^{nR_2}\}$.

Canais de Difusão (*Broadcast Channel*) são tipos de comunicação muito comuns e podem ser observados, por exemplo, pela analogia de um professor ministrando uma aula. Neste caso, há um transmissor (professor) e múltiplos receptores (alunos). Também podem ser observados em redes de computadores, transmissões de telefone, rádio, televisão entre outras TICs presentes no cotidiano (COVER; THOMAS, 2012).

No exemplo citado de sala de aula, há alunos que compreendem facilmente a informação e outros que tem dificuldades. Ao preparar uma aula, o método mais fácil a ser adotado pelo professor é usar o limite do aluno menos preparado em termos de instrumentos acadêmi-

[3] *"[...] an input alphabet X and two output alphabets, Y_1 and Y_2, and a probability transition function $p(y_1, y_2|X)$."*

cos, de modo de todos recebam o mínimo de informação necessária. No entanto, uma aula ideal deveria ser codificada e transmitida de modo que os alunos menos instrumentalizados recebessem o mínimo necessário e os alunos mais instrumentalizados recebessem uma codificação diferente, com mais informações. O mesmo exemplo se aplica a uma estação de televisão, cuja capacidade do canal pode estar limitada pelo pior receptor. No entanto, também podem ser codificadas de maneira diferente, de modo que a capacidade de canal ociosa dos melhores receptores (de som e imagem) seja aproveitada e estes possam receber imagem ou som de melhor qualidade. Desta forma, o objetivo passa a ser o de obter-se o mínimo quando o canal estiver ruim e obter um extra quando o canal estiver bom (COVER; THOMAS, 2012).

De acordo com Cover e Thomas (2012) e Gamal e Kim (2011), a forma de determinar a capacidade de canal nestes tipos de transmissões que utilizam múltiplos transmissores ou receptores é feita por regiões. No exemplo da sala de aula, considerando que cada aluno é um receptor diferente e que cada um pode receber informação também em condições diferentes (velocidade e alfabeto), as capacidades dos canais não seriam iguais, o que implicaria na necessidade de individualizar por aluno a codificação de fonte (compressão utilizada na informação transmitida). Também haveria regiões da sala onde o ruído seria maior e regiões onde o mesmo será menor. Desta forma, cada receptor necessitaria também de uma codificação específica de canal (para detecção e correção de erros). Por isso, a região de capacidade é o conjunto de taxas que podem ser alcançadas simultaneamente.

No caso da problemática específica para uma transmissão no formato *Broadcast Channel*, o desafio é o de encontrar taxas viáveis para transmissão de modo que se alcance a capacidade máxima a ser transmitida. Uma proposta é a de que a codificação seja usada para que os receptores disponibilizem as informações em camadas que se sobrepõem (COVER, 1975). No entanto, para Cover e Thomas (2012), ainda é distante a existência de uma teoria completa para esses casos e que por isso, ainda há uma série de problemas abertos na área. Mesmo se tal teoria for encontrada, sua implementação seria difícil devido à sua complexidade.

Por isso, ainda permeiam dúvidas: como codificar um sinal comum para múltiplos receptores? Como codificar e decodificar? Qual é a capacidade de canal máxima destas transmissões? Qual é a entropia? Gamal e Kim (2011), em seu livro, citam ao final de cada capítulo alguns problemas associados ao conceito de Teoria da Informação em Rede que precisam ser

melhor estudados. Dentre eles, pode-se destacar alguns relacionados a Canais de Difusão (*Broadcast Channel - BC*)[4]:

1. *"Capacity region of multiple access channels"* (GAMAL; KIM, 2011, p. 99);
2. *"What is the capacity region of less noisy BCs with four or more receivers?"* (GAMAL; KIM, 2011, p. 125);
3. *"Binary erasure broadcast channel"* (GAMAL; KIM, 2011, p. 126);
4. *"Product of two degraded broadcast channels"* (GAMAL; KIM, 2011, p. 126);
5. *"Product of two Gaussian broadcast channels"* (GAMAL; KIM, 2011, p. 127);
6. *"Reversely degraded broadcast channels with common message"* (GAMAL; KIM, 2011, p. 127);
7. *"Common-message broadcasting with state information"* (GAMAL; KIM, 2011, p. 195);
8. *"Separate source and channel coding over a DM-BC"* (GAMAL; KIM, 2011, p. 356);
9. *"Broadcasting over the relay channel"* (GAMAL; KIM, 2011, p. 421);
10. *"Common-message feedback capacity of broadcast channels e Broadcast channel with feedback"* (GAMAL; KIM, 2011, p. 457);
11. *"What is the ergodic capacity region of the Gaussian fading BC under fast fading when the channel gain information is available only at the decoders?"* (GAMAL; KIM, 2011, p. 598);

Desta forma, quando considera-se o crescimento do uso e dependência de equipamentos de comunicação que utilizam-se de transmissões *Broadcast Channel*, estes problemas em aberto tornam-se interessantes não somente para estudos científicos, mas principalmente para o desenvolvimento de novas tecnologias.

[4]Os problemas foram extraídos de um livro de 2011. Como as inovações nessa área crescem em ritmo acelerado, pode acontecer que eles não sejam mais considerados problemas abertos. Essa investigação, entre outras, relacionadas aos canais de difusão, é tema de trabalhos futuros.

Capítulo 4

Conclusões

Esta pesquisa realizou uma releitura de alguns conceitos essenciais de Teoria da Informação através de um resgate histórico, indicando um caminho onde há a descrição da gênese destes conceitos. O objetivo foi o de reduzir incertezas sobre alguns conceitos essenciais relacionados ao conceito de entropia em Teoria da Informação de modo que fosse possível indicar caminhos factíveis livres de interrogações e com abertura para desenvolvimentos tecnológicos e compreensão dos problemas, em teoria da informação, do mundo atual. Indicou-se também uma promissora via de estudo (uma estimativa para estudos futuros) na aplicação do conceito de entropia na resolução de problemas abertos em Teoria da Informação voltados à ideia de multicanais.

Nesse sentido buscou-se o esclarecimento de alguns termos na Teoria da Informação tais como: origem do nome entropia, "inventores" da teoria da informação e origem da letra H para denotar entropia.

1. Origem do termo "Teoria da Informação": apesar de Shannon não ter dado ênfase ao termo em seu artigo de 1948 e outros cientistas terem publicado antes de Shannon trabalhos em que constem este termo, já em 1945 há manuscritos que apontam sua utilização por Shannon. Nas pesquisas realizadas, não foi encontrado nenhum trabalho científico que usasse tal termo antes de 1945.

2. "Inventores" da Teoria da Informação: abordou-se a participação de diversos cientistas, Weaver, Wiener, Nyquist, Hartley, entre outros, no desenvolvimento das ideias voltadas à Teoria de Shannon. Notoriamente a Teoria da Informação foi concebida e desenvolvida através do acúmulo de conhecimento produzido ao longo dos anos que precederam o trabalho de Shannon. Por mais que outros trabalhos tenham sido desenvolvidos antes de Shannon (diga-se, de seu artigo seminal de 1948), não há como retirar dele o mérito de ter-se tornado uma "invenção" relevante para a história da humanidade. A criatividade no desenvolvimento de seu trabalho ao propor um modelo de comunicação fundamentado em distribuições estatísticas de um emissor a um receptor é um divisor de águas na história da ciência. Desse modo, mesmo tendo Wiener indicado modelos semelhantes ao

de Shannon, mas com propósitos diferentes, entende-se que deve-se atribuir a Shannon o mérito da invenção da teoria da informação tal como é entendida hoje em dia.

Um destaque especial foi dado ao conceito de entropia, que desde que foi apresentado por Clausius, abriu as portas do descobrimento científico e propiciou a Shannon a possibilidade de uma nova leitura para o conceito de entropia. A partir da termodinâmica, o conceito foi explorado, desenvolvido e expandido a outras áreas, originando interrogações que persistem até hoje. Assim como ocorreu na física, quando o termo foi usado por Shannon para batizar sua proposta matemática para medir a informação através da incerteza, naturalmente criaram-se novas interrogações. Dentro deste novo campo, a Teoria da Informação, este conceito tornou-se ainda mais plástico e complexo.

Dentre as interrogações criadas nesta dissertação, buscou-se apresentar argumentos para clarear a relação entre a entropia na Física e a entropia na Teoria da Informação, mostrando que apesar de utilizarem-se de fórmulas parecidas, e alguns autores conseguirem abstrair conceitos de igualdade como o de degradação da informação, as aplicações são muito diferentes, inclusive a simbologia. Porém, pode-se considerar que tanto os conceitos como a fórmula podem ter se originado na Física, sendo que o próprio Shannon reconheceu tal origem em seu trabalho. Há ainda "incertezas" com relação ao conceito de entropia que este trabalho investigou e não conseguiu clarear. Como, por exemplo, a escolha por Shannon do nome entropia para sua grandeza de medida de informação.

O fascínio por investigar as interrogações trazidas pelo conceito de entropia certamente foi e continuará sendo um excelente "combustível" utilizado para a investigação e desenvolvimentos de novas tecnologias. No passado, a entropia de Shannon foi o alicerce para a construção da capacidade de canal que, por sua vez possibilitou a construção das codificações de fonte e canal, eliminando problemas de redundância e erros, desenvolvendo as TICs como são conhecidas hoje. Por isso, a importância de Shannon para o desenvolvimento do conceito de entropia na Teoria da Informação é incontestável, porém as incertezas ainda existentes sobre esses conceitos (entropia na física e entropia na teoria da informação) ainda movem a ciência. Algumas destas incertezas podem ser clareadas com releituras sobre o conceito de entropia, assim como feito nesta dissertação. Há, porém, outras incertezas que surgem a partir destas releituras que apresentam-se como problemas abertos, como é o caso dos problemas relacionados a *Broadcast Channel*.

Por isso, para estudos futuros, considera-se a investigação do conceito de entropia associado ao problema de *Broadcast Channel*. Assim como a entropia foi o alicerce para a construção da Teoria da Informação, propõe-se que possa ainda estar no conceito de entropia a construção de uma forma otimizada de melhorar os canais de comunicação em *Broadcast*.

Considerando que a matemática foi a "matéria-prima" para construção de toda a Teoria da Informação, alicerçada no conceito de entropia, entende-se que ela seja essencial para o fortalecimento e expansão das TICs, devendo ser explorada e incentivada por aqueles que se predispõem levar à frente os desenvolvimentos científicos.

Por fim, espera-se que a contextualização, resgate histórico e clareamento de conceitos, realizados nesta dissertação, sejam reveladores da importância da Teoria da Informação no mundo moderno, e estimule estudantes, professores e pesquisadores a buscarem, com mais precisão, novos desenvolvimentos neste importante campo de pesquisa de modo que novas TICs possam também ser produzidas.

Referências Bibliográficas

AFTAB, O. *et al.* Information theory and the digital age. **Bandwagon**, p. 9–11, 2001.

AL-FEDAGHI, S. A conceptual foundation for the shannon-weaver model of communication. **International Journal of Soft Computing**, v. 7, n. 1, p. 12–19, 2012.

ALBINO, V.; GARAVELLI, A.; GORGOGLIONE, M. Organization and technology in knowledge transfer. **Benchmarking: An International Journal**, Emerald Group Publishing Limited, v. 11, n. 6, p. 584–600, 2004.

ASPRAY, W. F. The scientific conceptualization of information: A survey. **Annals of the History of Computing**, IEEE, v. 7, n. 2, p. 117–140, 1985.

BARNARD, G. The theory of information. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, v. 13, n. 1, p. 46–64, 1951.

BEECHER, M. D. Signalling systems for individual recognition: an information theory approach. **Animal Behaviour**, Elsevier, v. 38, n. 2, p. 248–261, 1989.

BEN-NAIM, A. **A Farewell to Entropy: Statistical Thermodynamics Based on Information : S = log W**. Danvers: World Scientific, 2008. ISBN 9789812707062.

BEN-NAIM, A. **Discover Entropy and the Second Law of Thermodynamics: A Playful Way of Discovering a Law of Nature**. Danvers: World Scientific, 2010. ISBN 9789814299756.

BEN-NAIM, A. **Information Theory: Part I: An Introduction to the Fundamental Concepts**. New Jersey: World Scientific, 2017. (Information Theory). ISBN 9789813208827.

BOLTZMANN, L. The second law of thermodynamics. In: ____. **Theoretical Physics and Philosophical Problems: Selected Writings**. Dordrecht: Springer Netherlands, 1974. p. 13–32. ISBN 978-94-010-2091-6.

BOLTZMANN, L. Further studies on the thermal equilibrium of gas molecules. **The Kinetic Theory Of Gases**, Cambridge, v. 1, p. 262–349, 2003. Série: History of Modern Physical Sciences, ISBN: 978-1-86094-347-8, Editado por Stephen G Brush e Nancy S Hall.

BOWN, R. Acoustics in communication. **The Journal of the Acoustical Society of America**, ASA, v. 21, n. 4, p. 305–307, 1949.

BRILLOUIN, L. Life, thermodynamics, and cybernetics. **American Scientist**, Sigma Xi, The Scientific Research Society, v. 37, n. 4, p. 554–568, 1949. ISSN 00030996.

BRILLOUIN, L. Thermodynamics and information theory. **American Scientist**, JSTOR, v. 38, n. 4, p. 594–599, 1950.

BRILLOUIN, L. **Science and Information Theory**. New York: Dover Publications, 1956.

BRISSAUD, J.-B. The meanings of entropy. **Entropy**, Molecular Diversity Preservation International, v. 7, n. 1, p. 68–96, 2005.

BROOKES, B. C. An introduction to the mathematical theory of information. **The Mathematical Gazette**, Mathematical Association, v. 40, n. 333, p. 170–180, 1956. ISSN 00255572.

BRUSH, S. Introduction. In: BRUSH, S. (Ed.). **Kinetic Theory**. Pergamon, 1966. p. 3 – 18. ISBN 978-0-08-011870-3. Disponível em: <http://www.sciencedirect.com/science/article/pii/B9780080118703500064>. Acesso em: 17 jan. 2019.

BURBURY, S. H. Xxxv. on some problems in the kinetic theory of gases. **Philosophical Magazine**, v. 30, n. 185, p. 298–317, 1890.

CAMPBELL, J. **Grammatical man: information, entropy, language, and life**. New York: Simon and Schuster, 1982. (Colección Popular). ISBN 9780671440619.

CAPES. **Portal de Periódicos - Busca pela palavra chave Shannon-Hartley**. 2017. Disponível em: <http://www.periodicos.capes.gov.br>. Acesso em: 17 nov. 2017.

CATTADORI, I. M.; HAYDON, D. T.; HUDSON, P. J. Parasites and climate synchronize red grouse populations. **Nature**, Nature Publishing Group, v. 433, n. 7027, p. 737, 2005.

CHAPMAN, S. Boltzmann's h-theorem. **nature**, Nature Publishing Group, v. 139, n. 3526, p. 931, 1937.

CHERRY, E. C. A history of the theory of information. **Proceedings of the IEE-Part III: Radio and Communication Engineering**, IET, v. 98, n. 55, p. 383–393, 1951.

CLAUSIUS, R. **Abhandlungen über die mechanische Wärmetheorie**. Braunschweig: F. Vieweg, 1864. (Abhandlungen über die mechanische Wärmetheorie, v. 1). Disponível em: <https://gallica.bnf.fr/ark:/12148/bpt6k95149n/f47.image>. Acesso em: 14 jan. 2019.

CLAUSIUS, R. **The Mechanical Theory of Heat: With Its Applications to the Steam-engine and to the Physical Properties of Bodies**. London: J. Van Voorst, 1867. Versão editada por Hirst, T.A. Disponível em: <https://books.google.com.br/books?id=8LIEAAAAYAAJ&printsec=frontcover&hl=pt-BR&source=gbs_ge_summary_r&cad=0#v=onepage&q=entropy&f=false>. Acesso em: 13 jan. 2019.

CONWAY, F.; SIEGELMAN, J. **Dark Hero of the Information Age: In Search of Norbert Wiener, The Father of Cybernetics**. New York: Basic Books, 2006. ISBN 9780465013715.

COOLEY, J. W.; TUKEY, J. W. An algorithm for the machine calculation of complex fourier series. **Mathematics of computation**, JSTOR, v. 19, n. 90, p. 297–301, 1965.

COVER, T.; THOMAS, J. **Elements of Information Theory**. New Jersey: Wiley, 2012. ISBN 9781118585771.

COVER, T. M. Open problems in information theory. In: IEEE PRESS. **1975 IEEE Joint Workshop on Information Theory**. New York, 1975. p. 35–36.

DIAS, P. M. C. A hipótese estatística do teorema-h. **Química Nova**, v. 17, n. 6, p. 472–479, 1994.

DIETZOLD, R. Network theory comes of age. **Electrical Engineering**, IEEE, v. 67, n. 9, p. 895–899, 1948.

DOOB, J. Review of ce shannon's the mathematical theory of communication. **Math. Rev.**, v. 10, p. 133, 1949.

DTU. **Welcome to the demonstration of RS en-/decoding in DVD players**. sem data. Disponível em: <http://www2.mat.dtu.dk/people/T.Hoeholdt/DVD/index.html>. Acesso em: 18 jan. 2019.

ECO, U. **A Theory of Semiotics**. Bloomington: Indiana University Press, 1976. (Advances in Semiotics). ISBN 9780253013316.

ECO, U. **A Estrutura Ausente: Introdução à Pesquisa Semiológica. 7 ed.,**. São Paulo: Perspectiva, 2001.

ELER, G. **Cientistas conseguem "ler a mente" do Demônio de Maxwell**. 2017. Disponível em: <https://super.abril.com.br/ciencia/cientistas-conseguem-ler-a-mente-do-demonio-de-maxwell>. Acesso em: 01 abr. 2018.

ELLERSICK, F. A conversation with claude shannon. **IEEE Communications Magazine**, v. 22, n. 5, p. 123–126, May 1984. ISSN 0163-6804.

FANO, R. M. **The transmission of information**. Massachusetts: Massachusetts Institute of Technology, Research Laboratory of Electronics Cambridge, 1949.

FANO, R. M. The information theory point of view in speech communication. **The Journal of the Acoustical Society of America**, ASA, v. 22, n. 6, p. 691–696, 1950.

FIORENTINI, D.; MIORIN, M. n.; MIGUEL, A. A contribuição para repensar... a educação algébrica elementar. **Pro-Posições**, v. 4, n. 1, p. 78–91, 2016. ISSN 1982-6248.

GALLAGER, R. G. Claude e. shannon: a retrospective on his life, work, and impact. **IEEE Transactions on Information Theory**, v. 47, n. 7, p. 2681–2695, 2001.

GAMAL, A. E.; KIM, Y.-H. **Network Information Theory**. New York: Cambridge University Press, 2011.

GAPPMAIR, W. Claude e. shannon: The 50th anniversary of information theory. **IEEE Communications Magazine**, IEEE, v. 37, n. 4, p. 102–105, 1999.

GOLDMAN, S. Some fundamental considerations concerning noise reduction and range in radar and communication. **Proceedings of the IRE**, IEEE, v. 36, n. 5, p. 584–594, 1948.

GOLDMAN, S. *et al.* **Communications and Related Projects**. Massachusetts, 1947.

GOLOMB, S. W. *et al.* Claude elwood shannon. **Notices of the AMS**, v. 49, n. 1, 2002.

GOOGLE. **Tradutor**. 2017. Disponível em: <https://translate.google.com.br/>. Acesso em: 20 nov. 2017.

GUIZZO, E. M. **The essential message: Claude Shannon and the making of information theory**. Tese (Doutorado) — Massachusetts Institute of Technology, 2003.

HAMMING, R. W. Error detecting and error correcting codes. **The Bell System Technical Journal**, v. 29, n. 2, p. 147–160, April 1950. ISSN 0005-8580.

HAMMING, R. W. **Art of Probability**. Redwood City: Addison Wesley Publishing Company, 1991.

HARTLEY, R. V. Transmission of information. **Bell Labs Technical Journal**, Wiley Online Library, v. 7, n. 3, p. 535–563, 1928.

HOSSENFELDER, S. **10 physics facts you should have learned in school but probably didn't**. 2018. Disponível em: <http://backreaction.blogspot.com/2018/07/10-physics-facts-you-should-have.html>. Acesso em: 25 mar. 2019.

HUFFMAN, D. A. A method for the construction of minimum-redundancy codes. **Proceedings of the IRE**, v. 40, n. 9, p. 1098–1101, Sept 1952. ISSN 0096-8390.

JAYNES, E. How does the brain do plausible reasoning? In: **Maximum-entropy and Bayesian methods in science and engineering**. Dordrecht: Springer, 1988. p. 1–24.

JAYNES, E. T. Information theory and statistical mechanics. **Phys. Rev.**, American Physical Society, v. 106, p. 620–630, May 1957.

JAYNES, E. T. Information theory and statistical mechanics. In: **Statistical Physics, 1962 Brandeis Lectures**. [S.l.]: New York, 1963. v. 3.

KINCHIN, A. I. **Mathematical theory of information**. New York: Dover Publications, 1957.

KULLBACK, S. Statistics and information theory. **J Wiley Sons, New York**, 1959.

LESNE, A. Shannon entropy: a rigorous mathematical notion at the crossroads between probability, information theory, dynamical systems and statistical physics. **Mathematical Structures in Computer Science**, p. 1–40, 2011.

LICKLIDER, J. The intelligibility of amplitude-dichotomized, time-quantized speech waves. **The Journal of the Acoustical Society of America**, ASA, v. 22, n. 6, p. 820–823, 1950.

LIN, S.; COSTELLO, D. **Error Control Coding: Fundamentals and Applications**. New Jersey: Prentice-Hall, 1983. (Prentice-Hall computer applications in electrical engineering series). ISBN 9780132837965.

LUNDHEIM, L. On shannon and"shannon's formula". **Telektronikk**, TELEDIREKTORATET, v. 98, n. 1, p. 20–29, 2002.

MAGOSSI, J. C.; PAVIOTTI, J. R. Incerteza em entropia. **Revista Brasileira de História da Ciência**. Submetido em 12/2018.

MANDREKAR, V.; MASANI, P. **Proceedings of the Norbert Wiener Centenary Congress, 1994: Michigan State University, November 27-December 3, 1994**. East Lansing: American Mathematical Soc., 1997. (AMS short course lecture notes). ISBN 9780821867570.

MASSEY, J. Information theory: The copernican system of communications. **IEEE Communications Magazine**, IEEE, v. 22, n. 12, p. 26–28, 1984.

MATWORKS. **Documentation - Huffman Coding**. Disponível em: <https://www.mathworks.com/help/comm/ug/huffman-coding-1.html>. Acesso em: 04 nov. 2017.

MATWORKS. **Documentation - huffmanenco**. Disponível em: <https://www.mathworks.com/help/comm/ref/huffmanenco.html>. Acesso em: 04 nov. 2017.

MAXWELL, J. **Theory of Heat**. [S.l.]: London, 1871.

MCELIECE, R. J. The reliability of computer memories. **Scientific American**, JSTOR, v. 252, n. 1, p. 88–95, 1985.

MCMILLAN, B. The basic theorems of information theory. **The Annals of Mathematical Statistics**, JSTOR, p. 196–219, 1953.

MILLER, G. A. Language engineering. **The Journal of the Acoustical Society of America**, ASA, v. 22, n. 6, p. 720–725, 1950.

MIT. **MIT Professor Claude Shannon dies; was founder of digital communications**. 2001. Disponível em: <http://news.mit.edu/2001/shannon>. Acesso em: 01 out. 2017.

MOON, T. **Error Correction Coding: Mathematical Methods and Algorithms**. New Jersey: Wiley, 2005. ISBN 9780471739142.

MOREIRA, M. M.; PICADO, J. Truques e magia com códigos algébricos. **Gazeta de Matemática**, n. 175, p. 20–29, 03 2015. Disponível em: <http://gazeta.spm.pt/getArtigo?gid=481>. Acesso em: 17 jan. 2019.

MOSER, S.; CHEN, P. **A Student's Guide to Coding and Information Theory**. New York: Cambridge University Press, 2012. ISBN 9781107601963.

MOUILLOT, D.; LEPRETRE, A. A comparison of species diversity estimators. **Researches on Population Ecology**, Springer, v. 41, n. 2, p. 203–215, 1999.

MOUTON, A. M. *et al.* Optimisation of a fuzzy physical habitat model for spawning european grayling (thymallus thymallus l.) in the aare river (thun, switzerland). **ecological modelling**, Elsevier, v. 215, n. 1, p. 122–132, 2008.

NALON, J. A. **Introdução ao processamento digital de sinais**. Rio de Janeiro: LTC, 2009. ISBN 9788521616467.

NEBEKER, F. Birth certificate of the information age: The annus mirabilis 1948. In: SIGNAL PROCESSING: THE EMERGENCE OF A DISCIPLINE, 1948-1998. **IEEE Press**. New Jersey, 1998. p. 13–27.

NEBEKER, N. Fifty years of signal processing: The ieee signal processing society and its technologies 1948-1998. **The IEEE Signal Processing Society**, 1998.

NEUMANN, J. V. Proposal and analysis of a new numerical method for the treatment of hydrodynamical shock problems. **Collected works**, v. 6, p. 351–379, 1944.

NYQUIST, H. Certain factors affecting telegraph speed. **Transactions of the American Institute of Electrical Engineers**, XLIII, p. 412–422, Jan 1924. ISSN 0096-3860.

OLIVER, B.; PIERCE, J.; SHANNON, C. E. The philosophy of pcm. **Proceedings of the IRE**, IEEE, v. 36, n. 11, p. 1324–1331, 1948.

PAK, A.; PAROUBEK, P. Twitter as a corpus for sentiment analysis and opinion mining. In: **LREc**. [S.l.: s.n.], 2010. v. 10, n. 2010.

PAVIOTTI, J. R.; MAGOSSI, J. C. Considerações sobre o conceito de entropia na teoria da informação. In: **Anais do X WORKSHOP DA PÓS-GRADUAÇÃO DA FACULDADE DE TECNOLOGIA**. FT/UNICAMP - Limeira, 2018. p. 24. Disponível em: <https://wordpress.ft.unicamp.br/workshoppos/edicao-2018/>. Acesso em: 31 mar. 2019.

PIERCE, J. The early days of information theory. **IEEE Transactions on Information Theory**, v. 19, n. 1, p. 3–8, 1973.

PIERCE, J. **An Introduction to Information Theory: Symbols, Signals and Noise**. New York: Dover Publications, 2012. (Dover Books on Mathematics). 2ª Edição Revisada do Original publicado em 1961 por Harper Brohers. ISBN 9780486134970.

PIERCE, J. R. Men, machines, and languages. **IEEE spectrum**, IEEE, v. 5, n. 7, p. 44–49, 1968.

PINEDA, J. O. d. C. **A entropia segundo Claude Shannon: o desenvolvimento do conceito fundamental da teoria da informação. 2006. 124 f**. Tese (Doutorado) — Dissertação (mestrado em História da Ciência)-Pontifícia Universidade Católica de São Paulo, São Paulo, 2006.

REICH, E. **The Theory of Information**. Santa Mônica, 1950. Disponível em: <https://apps.dtic.mil/dtic/tr/fulltext/u2/116555.pdf>. Acesso em: 13 jan. 2019.

REZA, F. **An Introduction to Information Theory**. New York: Dover, 1961. (Dover Books on Mathematics Series). ISBN 9780486682105.

RIOUL, O.; MAGOSSI, J. C. On shannon's formula and hartley's rule: Beyond the mathematical coincidence. **Entropy**, v. 16, n. 9, p. 4892–4910, 2014. ISSN 1099-4300.

ROBICHAUD, M.-A. **File Exchange - Hamming Code**. 2013. Disponível em: <https://www.mathworks.com/matlabcentral/fileexchange/40208-hamming-code>. Acesso em: 04 nov. 2017.

SALOMON, D. **Data Compression: The Complete Reference**. Berlin: Springer Berlin Heidelberg, 2012. ISBN 9783642860928.

SEISING, R. On two 60 years old theories and the theory of fuzzy sets and systems: Cybernetics and information theory. In: **NAFIPS 2009 - 2009 Annual Meeting of the North American Fuzzy Information Processing Society**. [S.l.]: IEEE, 2009. p. 1–6.

SHANNON, C.; WEAVER, W. **The Mathematical Theory of Communication**. Urbana: University of Illinois Press, 1964.

SHANNON, C. E. A mathematical theory of cryptography. **Bell Tel. Lab. memo**, 1945. Posteriormente revisado e publicado como "Communication theory of secrecy systems," Bell Syst. Tech. J., vol. 28, pp. 656-715, Oct. 1949.

SHANNON, C. E. A mathematical theory of communication. **The Bell System Technical Journal**, v. 27, n. 3, p. 379–423, July 1948. ISSN 0005-8580.

SHANNON, C. E. Communication theory of secrecy systems. **Bell Labs Technical Journal**, Wiley Online Library, v. 28, n. 4, p. 656–715, 1949.

SHANNON, C. E. The bandwagon. **IRE Transactions on Information Theory**, v. 2, n. 1, p. 3, 1956.

SHPAK, M.; CHURCHILL, G. A. The information content of a character under a markov model of evolution. **Molecular phylogenetics and evolution**, Elsevier, v. 17, n. 2, p. 231–243, 2000.

SZILARD, L. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. **Systems Research and Behavioral Science**, Wiley Online Library, v. 9, n. 4, p. 301–310, 1964.

TAKADA, H. H. On execution strategies and minimum discrimination information principle. In: AIP PUBLISHING. **AIP Conference Proceedings**. [S.l.], 2016. v. 1757, n. 1, p. 050003.

TAKADA, H. H.; SANTOS, R. A. Portfolio diversification using information theory applied to brazilian stocks. **Journal of Mathematics and System Science**, David Publishing Company, Inc., v. 4, n. 5, 2014.

TEMPLET, P. H. Energy, diversity and development in economic systems; an empirical analysis. **Ecological Economics**, Elsevier, v. 30, n. 2, p. 223–233, 1999.

TOEPLITZ, O. **The calculus: a genetic approach**. Chicago: University of Chicago Press, 2008.

TRIBUS, M.; MCIRVINE, E. C. Energy and information. **Scientific American**, Scientific American, a division of Nature America, Inc., v. 225, n. 3, p. 179–190, 1971. ISSN 00368733, 19467087.

VERDU, S. Fifty years of shannon theory. **IEEE Transactions on information theory**, IEEE, v. 44, n. 6, p. 2057–2078, 1998.

VERDÚ, S.; MCLAUGHLIN, S.; SOCIETY, I. I. T. **Information Theory: 50 Years of Discovery**. [S.l.]: Wiley, 2000. ISBN 9780780353633.

VILLANI, C. **H-Theorem and beyond: Boltzmann's entropy in today's mathematics**. 2008. Disponível em: <http://docplayer.net/30183507-H-theorem-andbeyond-boltzmann-s-entropy-in-today-s-mathematics.html>. Acesso em: 13 jan. 2019.

WANG, N.; LI, B. Proceedings a model of deceitful information communication: Some views on theory and practice of semantic information. In: **Multidisciplinary Digital Publishing Institute Proceedings**. [S.l.: s.n.], 2017. v. 1, n. 3, p. 127.

WEAVER, W. The mathematics of communications. **Scientific American**, Scientific American, v. 181, n. 1, p. 11–15, 1949.

WEHRL, A. General properties of entropy. **Rev. Mod. Phys.**, American Physical Society, v. 50, p. 221–260, Apr 1978.

WIENER, N. Cybernetics. **Cambridge, Mass**, 1948.

WIENER, N. What is information theory. **IRE Transactions on Information Theory**, v. 2, n. 2, p. 48, 1956.

WIENER, N. **Cybernetics Or Control and Communication in the Animal and the Machine**. Cambridge: M.I.T. Press, 1961. (DE-601)251474038: MIT paperback series). ISBN 9780262730099.

WIENER, N. **Cibernética e sociedade: o uso humano de seres humanos**. São Paulo: Cultrix, 1973.

WOODWARD, P. M.; DAVIES, I. L. Information theory and inverse probability in telecommunication. **Proceedings of the IEE-Part III: Radio and Communication Engineering**, IET, v. 99, n. 58, p. 37–44, 1952.

Apêndice A

Exemplos de Codificação de Huffmann em Matlab

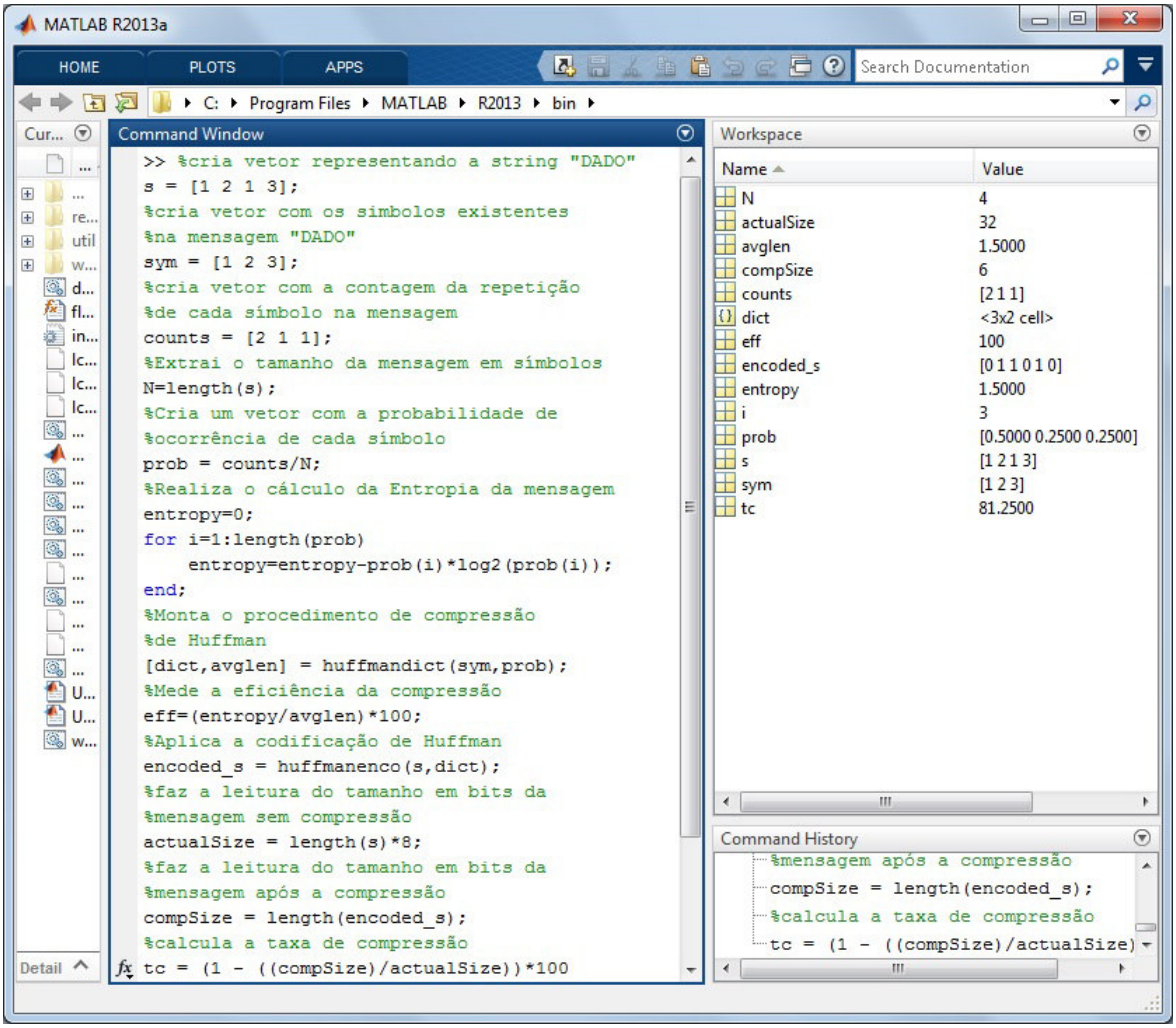
Os algoritmos a seguir foram usados para testar e validar os exemplos apresentados no capítulo 3 desta dissertação através do *software MATLAB*. Tais algoritmos foram extraídos e adaptados do *site* do fabricante do *MATLAB*, *Matworks*. Nesta primeira representação, os conjuntos de símbolos M_1 e M_2 foram representandos através de um vetor s de números inteiros, pois as funções que realizam a codificação de Huffman da ferramenta *MATLAB* não são compatíveis com símbolos representados por caracteres alfanuméricos (MATWORKS, a; MATWORKS, b).

```
%cria vetor representando a string "DADO"
s = [1 2 1 3];
%cria vetor com os simbolos existentes na mensagem "DADO"
sym = [1 2 3];
%cria vetor com a contagem da repeticao de cada simbolo na
   mensagem
counts = [2 1 1];
%Extrai o tamanho da mensagem em simbolos
N=length(s);
%Cria um vetor com a probabilidade de ocorrencia de cada
   simbolo
prob = counts/N;
%Realiza o calculo da Entropia da mensagem
entropy=0;
for i=1:length(prob)
    entropy=entropy-prob(i)*log2(prob(i));
end;
%Monta o procedimento de compressao de Huffman
[dict,avglen] = huffmandict(sym,prob);
%Mede a eficiencia da compressao
eff=(entropy/avglen)*100;
%Aplica a codificacao de Huffman
```

```
encoded_s = huffmanenco(s,dict);
%faz a leitura do tamanho em bits da mensagem sem compressao
actualSize = length(s)*8;
%faz a leitura do tamanho em bits da mensagem apos a compressao
compSize = length(encoded_s);
%calcula a taxa de compressao
tc = (1 - ((compSize)/actualSize))*100
```

A Figura A.1 o algoritmo utilizado, onde o conjunto de símbolos $M_1 = [D,A,D,O]$ foi representado através do um vetor s = [1 2 1 3]. Como observa-se, o cálculo da eficiência e da taxa de compressão através do *MATLAB* apresentaram respectivamente 100% e 81,25%.

Figura A.1 - Cálculo da eficiência e taxa de compressão através do *MATLAB* - Exemplo 1



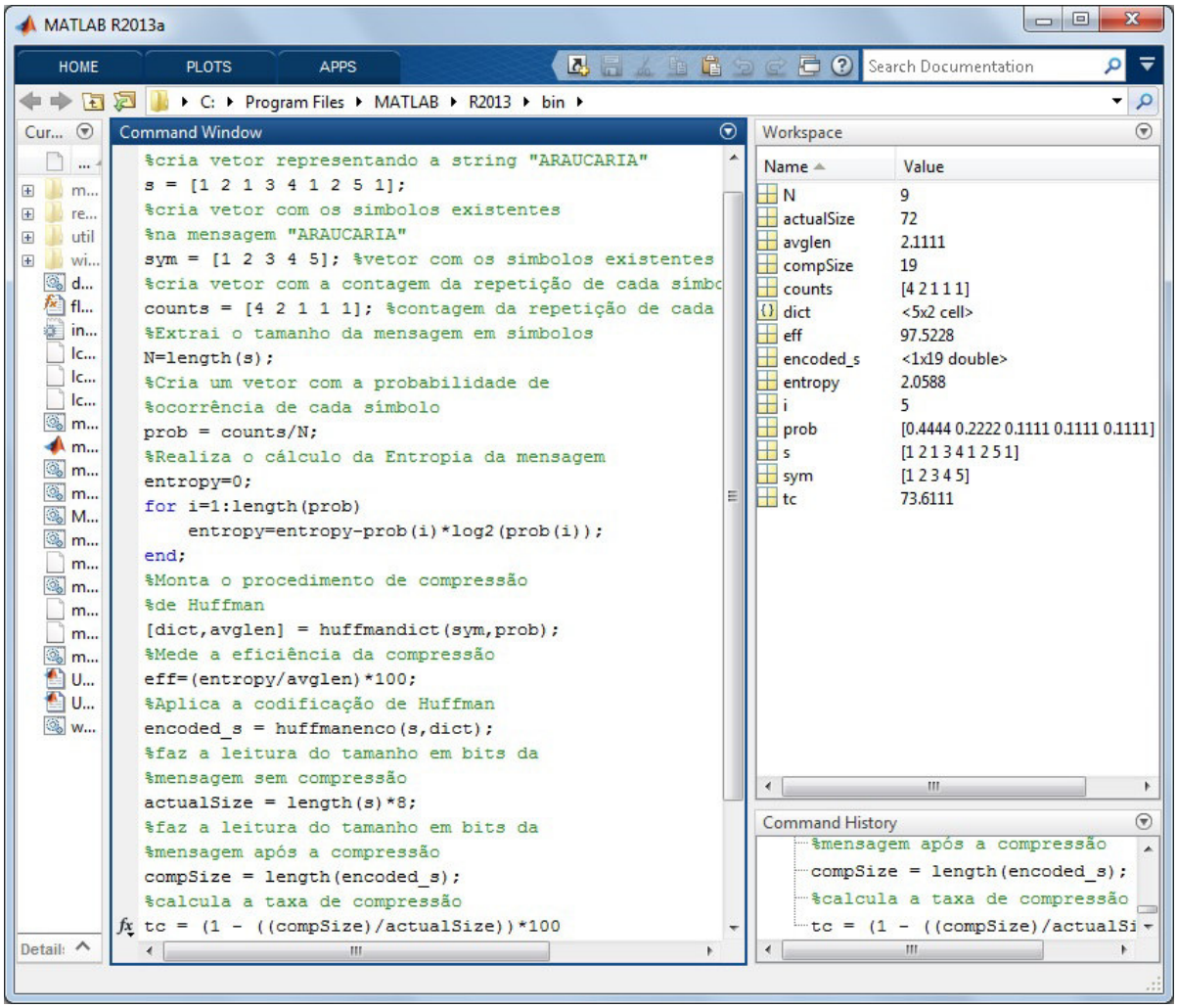
Fonte: Produção do próprio autor.

Os cálculos para o segundo exemplo também foram realizados através da ferramenta *MATLAB*. São demonstrados através algoritmo seguinte e da Figura A.2. Neste segundo exemplo, o conjunto $M_2 = [A,R,A,U,C,A,R,I,A]$ foi representado através do vetor s = [1 2 1 3 4 1 2 5 1]. Neste outro exemplo, observa-se a obtenção da eficiência de compressão de 97,52% e a taxa de compressão de 76,61%.

```
%cria vetor representando a string "ARAUCARIA"
s = [1 2 1 3 4 1 2 5 1];
%cria vetor com os simbolos existentes na mensagem "ARAUCARIA"
sym = [1 2 3 4 5]; %vetor com os simbolos existentes na
   mensagem
%cria vetor com a contagem da repeti  de cada simbolo
counts = [4 2 1 1 1]; %contagem da repeticao de cada simbolo na
    mensagem
%Extrai o tamanho da mensagem em simbolos
N=length(s);
%Cria um vetor com a probabilidade de ocorrencia de cada
   simbolo
prob = counts/N;
%Realiza o calculo da Entropia da mensagem
entropy=0;
for i=1:length(prob)
    entropy=entropy-prob(i)*log2(prob(i));
end;
%Monta o procedimento de compressao de Huffman
[dict,avglen] = huffmandict(sym,prob);
%Mede a eficiencia da compressao
eff=(entropy/avglen)*100;
%Aplica a codificacao de Huffman
encoded_s = huffmanenco(s,dict);
%faz a leitura do tamanho em bits da mensagem sem compressao
actualSize = length(s)*8;
%faz a leitura do tamanho em bits da mensagem apos a compressao
```

```
compSize = length(encoded_s);
%calcula a taxa de compressao
tc = (1 - ((compSize)/actualSize))*100
```

Figura A.2 - Cálculo da eficiência e taxa de compressão através do *MATLAB* - Exemplo 2



Fonte: Produção do próprio autor.

Apêndice B

Exemplos de Codificação de Hamming em Matlab

Um dos exemplos de código de Hamming abordado na seção sobre Codificação de Canal, no capítulo 3 desta dissertação, também foi testado e validado através de algoritmo no *software MATLAB*. Este algoritmo é apresentado a seguir. Foi adaptado de Robichaud (2013), disponibilizado no fabricante do *MATLAB*.

```
%Define a variavel m (mensagem a ser codificada)
m=[0 1 0 1 1];
%Verifica o tamanho de m em quantidade de digitos
tam=length(m);
%Atraves do tamanho, define a quantidade de digitos necessarios
    para a checagem de paridade (k)
k=floor(log2(tam+ceil(log2(tam))))+1;
%cria a variavel n para receber a codificacao do tamnaho de m +
    k
n=ones(1,length(m)+k);
%insere zeros nas posicoes de k (2^n) em n
for I=0:k
  n(1,2^I)=0;
  n=n(1,1:length(m)+k);
end
%insere os digitos de m nas demais posicoes de n
for J=1:length(n)
        if n(1,J)==1
                count=floor(log2(J)+1);
        n(1,J)=m(1,J-count);
        end
end
%cria matriz (tabela verdade) para gerar digitos de paridade k
P=zeros(k,length(n));
stop_z=length(P);
```

```
for X=1:k
        for Y=0:length(P)-1
                if Y<stop_z/2^X
                        P(X,((2^X)*Y+2^(X-1)):((2^X)*Y+(2^X)-1)
                           )=1;
                end
        end
end
P=P(:,1:stop_z);
for V=1:k
    Q(V,:)=P(V,:).*n;
end
for U=1:k
    R(U,:)=mod(length(find(Q(U,:))),2);
end
for S=0:k-1
    n(1,2^S)=R(S+1,1);
end
disp(n);
```

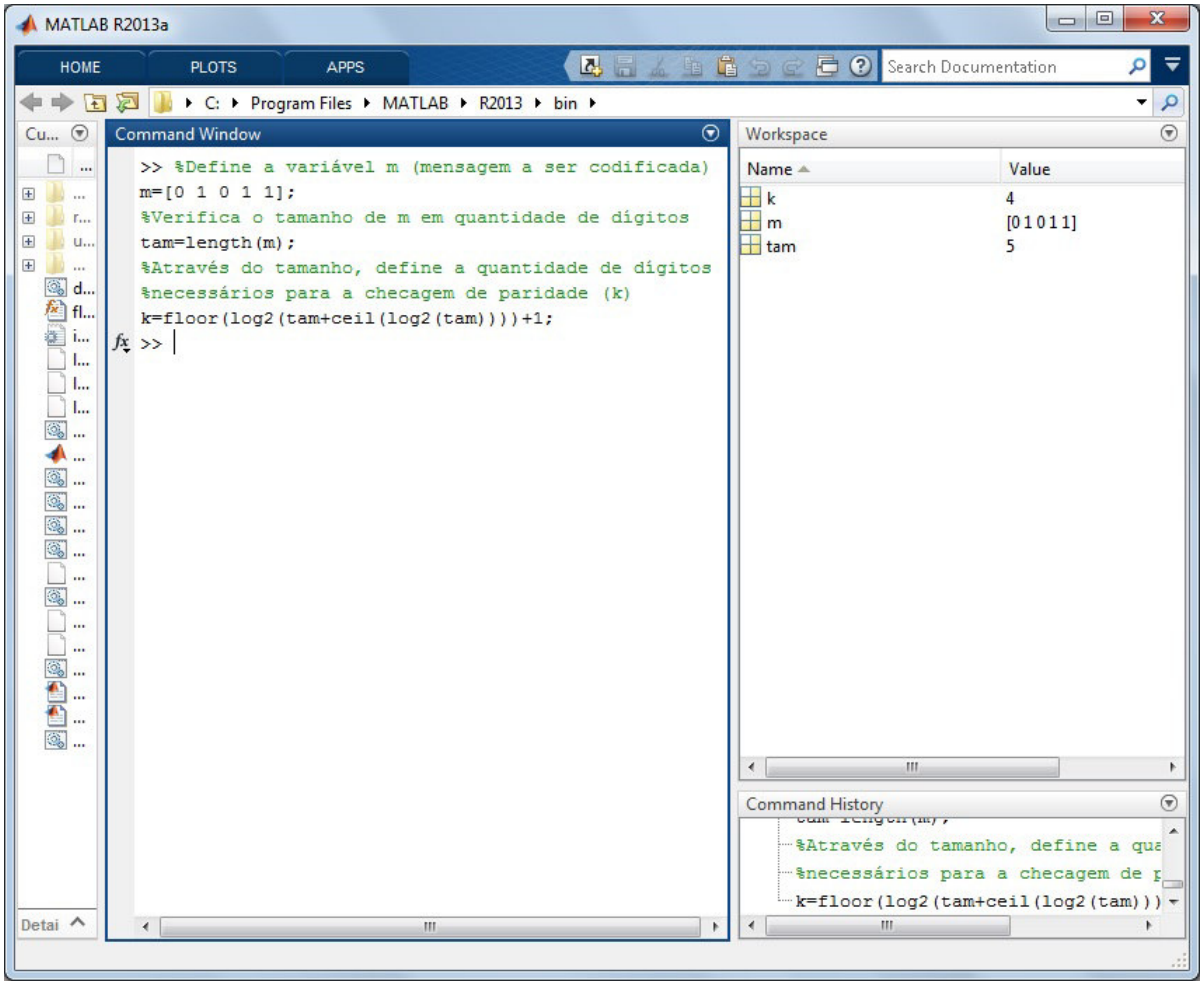
Para facilitar a compreensão, as Figuras B.1, B.2 e B.3 representam em etapas o processo de codificação de Hamming.

1. Verifica-se o tamanho de m em quantidade de dígitos para se determinar quantos dígitos de checagem de paridade k serão adicionados para a codificação de m em n. A Figura B.1 representa esta primeira etapa sendo reproduzida no MATLAB.

2. Sabendo que serão necessários 4 dígitos de checagem paridade, então sabe-se também o tamanho da mensagem codificada n, pois $n = m + k$, portanto $n = 9$. Os 4 dígitos de checagem, ocuparão as posições 1, 2, 4 e 8 de n. Tais procedimentos são reproduzidos em MATLAB e apresentados na Figura B.2.

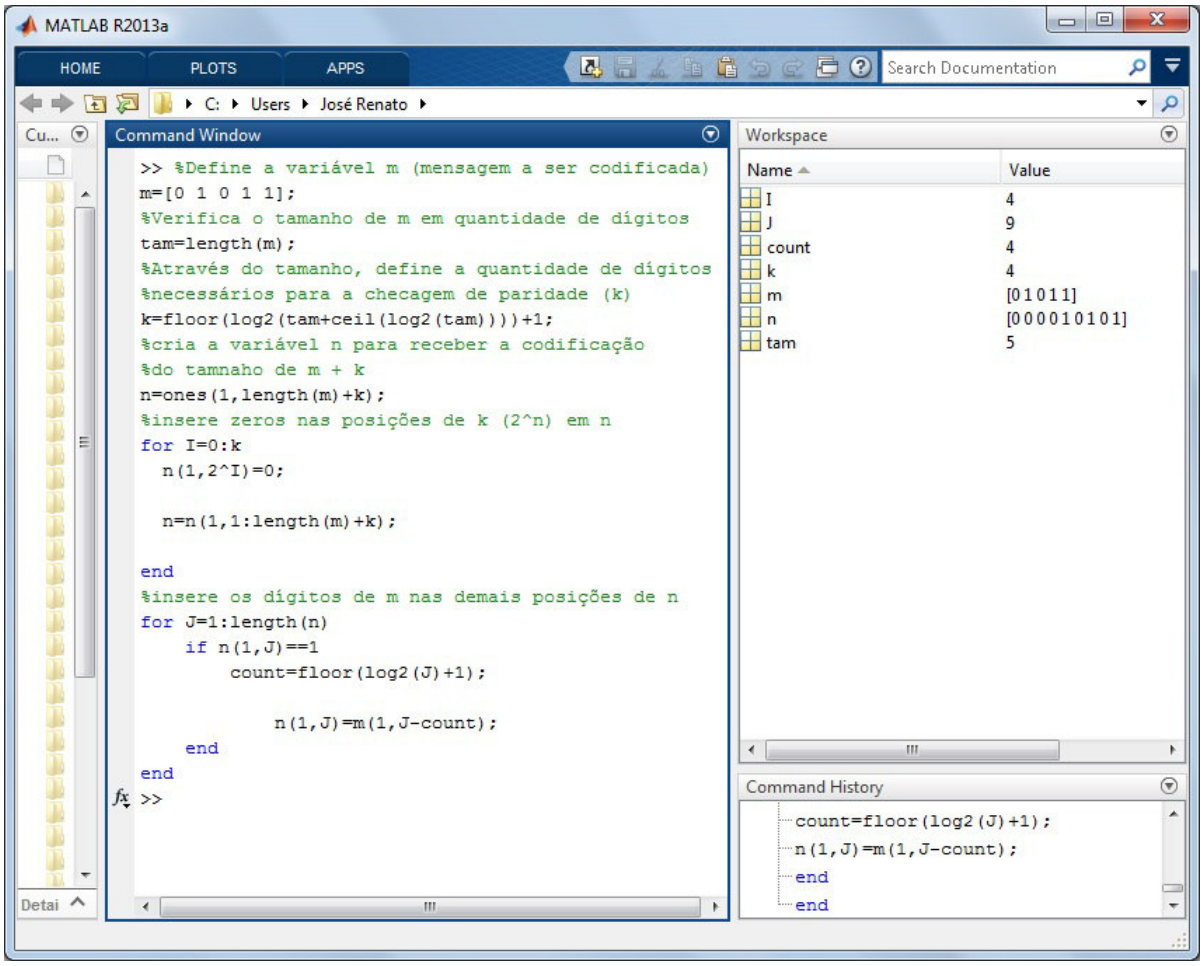
3. Determina-se a o valor de cada dígito de paridade. Tal procedimento é reproduzido na Figura B.3.

Figura B.1 - Codificação de Hamming - Etapa 1 em MATLAB



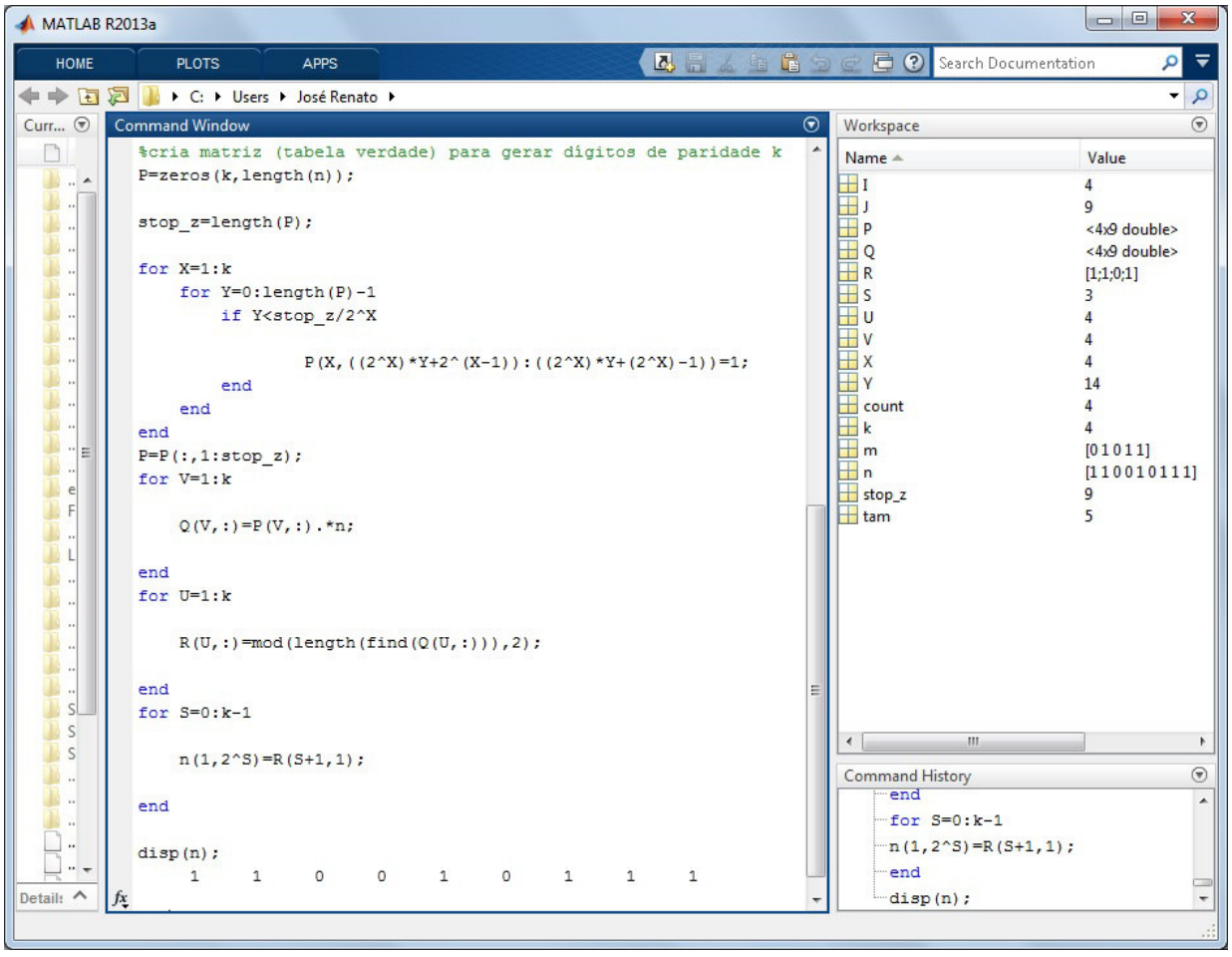
Fonte: Produção do próprio autor.

Figura B.2 - Codificação de Hamming - Etapa 2 em MATLAB



Fonte: Produção do próprio autor.

Figura B.3 - Codificação de Hamming - Etapa 3 em MATLAB



Fonte: Produção do próprio autor.